Genetic dissection of secondary metabolite production during heartwood formation of Western redcedar (*Thuja plicata*)

by

Sifat Tasnim

B.Sc. (Biotechnology & Genetic Engineering), Khulna University, 2014

Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of
Doctor of Philosophy

in the
Department of Biological Sciences
Faculty of Science

SIMON FRASER UNIVERSITY
Fall 2023

Declaration of Committee

Name:	**Sifat Tasnim**
Degree:	**Doctor of Philosophy**
Title:	**Genetic dissection of secondary metabolite production during heartwood formation on Western redcedar (*Thuja plicata*)**
Committee:	**Chair:** **Norbert Haunerland** Professor, Biological Sciences
	Jim Mattsson Supervisor Professor, Biological Sciences
	Jennifer Cory Committee Member Professor, Biological Sciences
	Sherryl Bisgrove Committee Member Associate Professor, Biological Sciences
	Harald Hutter Examiner Professor, Biological Sciences
	Dezene Huber External Examiner Professor, Biochemistry & Molecular Biology University of Northern British Columbia

Abstract

Thuja plicata, commonly known as Western redcedar, is an economically important conifer species on the North-west coast of Canada and the USA. *T. plicata* lumber is valued for its natural durability. Paradoxically, logged trees are frequently culled because of extensive heartwood (HW) fungal rot, indicating genetic variation in rot resistance. Resistance to rot in the dead HW of trees is attributed primarily to secondary metabolites. *T. plicata* HW has high levels of fungistatic lignans and fungitoxic monoterpene-derived tropolones, primarily β-thujaplicin. However, it takes 15-20 years of growth before β-thujaplicin is produced at high levels, precluding breeding based on β-thujaplicin levels. If critical genes behind terpenoid biosynthesis were known, variants could be correlated with high levels of corresponding terpenoid, allowing selection for HW rot resistance based on genotype of seedlings after crosses. In a quest for relevant genes, we used RNA-seq to identify genes upregulated in the sap to heart wood transition zone in which secondary metabolites are produced. Among them, we found six putative terpene synthases (TpTSs) and three putative 2-oxoglutarate/Fe(II)-dependent dioxygenases (TpDOXs) that are candidates for synthesis of specific terpenes and for ring expansion of monoterpenes into tropolones, respectively. Enzyme assays using recombinant proteins and terpene substrates showed that one enzyme produced the monoterpene terpinolene and another terpinolene and 3-carene, both of which may be precursors to β-thujaplicin. One terpene synthase produced the sesquiterpene alcohol elemol, previously linked to termite resistance. In addition, we identified three monofunctional diterpene synthases, of which one produced the intermediate normal-copalyl diphosphate, an intermediate for diterpene biosynthesis, the second produced sandaracopimaradiene or *syn*-stemod-13(17)-ene depending on the substrate, and the third produced levopimaradiene. Enzyme assay of three TpDOXs with a potential substrate did not yield any β-thujaplicin. Taken together, we have identified genes and enzymes that for the first time provides detailed insights into terpene biosynthesis during heartwood formation in *T. plicata*. Lastly, we show that RNA-seq methodology and identified genes can be used also for field studies of biotic and abiotic stress in *T. plicata* silviculture.

Keywords: terpene synthase; β-thujaplicin; heartwood; secondary metabolites; western redcedar; Fe(II)/2-oxoglutarate dependent dioxygenase

Dedication

My dear son FAYAZ, who I am dedicating this thesis to. Thank you for teaching me how to be stronger, patient, and wiser in my life. Your mother loves you to the moon and back.

Acknowledgements

First and foremost, I would like to thank my mentor, Dr. Jim Mattsson for the continued assistance and advice that he has given me during this period of study and research. It was through his amazing guidance and patience that the work in this thesis could be carried out. He will be a source of my gratitude all the time.

I would like to thank Dr. Jennifer Cory, and Dr. Sherryl Bisgrove, who contributed considerable time and expertise as a member of my supervisory committee for their constructive input during our meetings. I would like to take this opportunity to say thank you very much.

Regine Gries has worked very hard and gave valuable advice on this project. Words can not express my gratitude to her in any way. Thanks also to Drs. Christopher Keeling, Lina Madilao, Joerg Bohlmann, Meimei Xu, Reuben J. Peters, Russell Cox, Priyadarshini Balaraman, Erika Plettner, Mike Cruikshank, David Dun, Gerhard Gries, and Tyler J. Doyon for their help in providing valuable resources and advice.

I would like to extend a special thanks to Don Wiggins, who assisted me with these studies and provided information that was useful for my work.

I am indebted to my current labmates, Sachi, Adam, Juan and the former labmates Chen, Tiffany, Nicole, Don, Stacey and Katherine, with whom I was privileged to share a laboratory. They have given me all sorts of assistance. I will never forget the encouragement and friendship you gave me.

A lot of thanks to my family for all their support and motivation over the course of this study. Without you, I would not have been able to make it.

This work is dedicated to my son, but I also want to devote it to each of the people mentioned above.

Table of Contents

List of Figures

List of Tables

Chapter 1. Introduction

1.1. Thuja plicata

1.1.1. Habitat & Distribution

Western redcedar (WRC, *Thuja plicata* Donn ex D. Don), also known as pacific redcedar, giant cedar, canoe cedar, or giant arborvitae, is one of the most resilient and versatile conifers that can live for over 2000 years [1]. These long-lived trees can grow up to 70 meters in height and 7 meters in diameter [2]. This species belongs to the Cupressaceae family of conifers, native to the Pacific coast of North America. In general, they are a shade-tolerant species with a shallow, wide-spreading root system and grow commonly in mixed-species stands [3,4]. The natural range of *T. plicata* is divided into two distinct regions, one coastal population growing from northern California to southeastern Alaska and a second interior population ranging from eastern Montana, and northern Idaho to eastern British Columbia [3]. In British Columbia (BC), it grows at low to mid-elevations in wet coastal and interior areas with mild temperatures. [5]. The cultural and ecological significance made *T. plicata* an official tree of BC on February 18, 1988 [6].

1.1.2. Economy & Uses

T. plicata wood products have high value, selling at two to four times higher prices than other commercially important North American species [7]. In BC, *T. plicata* lumber is among the highest-priced softwoods, generating more than $1 billion in sales annually [8]. *T. plicata* wood is durable, swells marginally upon wetting, and is easy to work. Given their natural durability, *T. plicata* wood products are used for exterior residential applications such as roof and wall coverings, fencing, decking, garden furniture, and planters [3,5]. *T. plicata* wood products are more environmentally friendly compared to those treated with biocides. In addition, the essential oil, rich in terpenoids, is used in the cosmetics and fragrance industry due to its antimicrobial and scent qualities. Among indigenous peoples in BC, this “Tree of Life” has for thousands of years been of special spiritual and practical importance, used for daily essentials such as hats, baskets, boxes, ropes, clothing, blankets, canoes, totem poles, and more [9]. The quality and economic value of *T. plicata* wood lies with its natural resistance against rot and

fungus. This resistance is attributed to the secondary metabolites also known as specialized metabolites or extractives produced by the heartwood of mature trees [10].

1.2. Heartwood

1.2.1. Heartwood formation process

The wood of gymnosperm and angiosperm tree species can be separated into two different regions based on function. The outer and younger sapwood (SW) region transports water through tracheary elements (TEs), whereas the inner and older heartwood (HW), does not transport water due to TEs being blocked by air embolisms and in many species also by secondary metabolites. The metabolites, in turn, provide resistance to HW degradation by microbes and insects, and typically also provide a darker color to the HW relative to the pale-colored SW [11] (Figure 1.1). The TEs are characterized by elongated and interconnected cells to facilitate water transport, and thick secondary walls to resist the negative pressure or tension caused by transpiration. In addition, wood also contains a fraction (5-25%) of thin-walled parenchyma cells in the form of radial bundles known as rays, and axial parenchyma cells [12,13]. In the transition zone (TZ) between SW and HW, the starch is used to produce and secrete secondary metabolites before the parenchyma cells die, resulting in HW in which all cells typically are dead [14–16]. In addition to developmental regulation, secondary metabolite production can also be triggered prematurely and closer to the bark in response to wounding and pathogen attack, sealing off healthy SW from the affected site [17,18].

The process of HW formation is still not well understood and is known primarily from physiological and anatomical studies on a small number of species [14]. Gene expression during heartwood formation has been studied to a limited extent, primarily in scots pine (*Pinus sylvestris* L.) & sandalwood (*Santalum album* L.) [11,19]. Based on recent studies, the above paradigm of heartwood formation has been refined into four different models of HW formation. In Robinia-type trees, secondary metabolites are produced only in TZ and absent in SW [19–21]. In the Juglans-type, low levels of secondary metabolism and accumulation occurs in SW, followed by a spike in the TZ [23]. *P. sylvestris* also falls into this category, as some metabolites like diterpenoid resin acids are produced in the SW but most, including, stilbenes are produced in the TZ [19],

results in part supported by a recent transcriptomics study. In the Taiwania-type, the biosynthesis of secondary metabolites occurs also in SW [24,25]. Finally, in the Santalum-type, some parenchyma cells remain live and active in the HW. In this model, preferentially, precursors for mono and diterpene biosynthesis are expressed in SW while genes for sesquiterpene biosynthesis are expressed in HW [11], correlating with the sesquiterpene alcohols santalols and bergamotol, well-known components of sandalwood-based perfumes [26,27].

Due to discrepancies between species and growing places, the exact season of HW formation is poorly understood [21]. Most data suggest that HW forms in the dormant season [15,16,28].

Figure 1.1. **Cross section of *T. plicata* stem.**

Bark: the outermost layer. Cambium: a layer of thin cells followed by bark. Pith: the centre of the trunk. Heartwood: darker coloured within the centre. Sapwood: lighter coloured on the outside. Transition zone: transition from sapwood to heartwood.

1.2.2. Heartwood rot

Rot or decay that develops primarily in the HW of stem is called heartwood rot. This rot is caused by a successive invasion of fungal agents and pathogens that can enter the core of the tree through open wounds, bark, or root system [7]. The fungi detoxify the HW substances and feed on and degrade wood components [29]. HW rot is a threat to the tree as it loses strength over time [30]. In a tropical forest, a variety of HW-attacking fungi and insects produce hollow cores that can affect over 30% of trees, including

softwoods and hardwoods [31]. In *T. plicata*, up to 30% of harvested log volumes are culled due to extensive HW rot [2]. More than 30 different decay-causing fungi have been shown to be associated with HW rot in *T. plicata* trees [2]. Studies have shown that two fungal species, *Armillaria ostoyae* and *Coniferiporia weirii (also known as Phellinus weirii* or *Poria weirii*) are the major causes of butt, root and heart rot in *T. plicata* trees, affecting up to 10% volume of coastal and 28% of interior populations [7,32,33]. Other organisms such as termites and bacteria can cause significant damage to *T.plicata* wood. In trembling aspen (*Populus tremuloides* Michx.), HW rot caused by *Phellinus tremulae* is the primary cause of volume loss in this species [34]. Moreover, HW rot in eastern white pine (*Pinus strobus* L.) causes significant volume loss, which reduces the price of lumber by 50% [35]. In studies of Norway spruce (*Picea abies* L. H. Karst.), 18-58% of 81-120 year-old trees were damaged by butt and root rot caused by *Heterobasidion annosum* [36–38]. The primary defense against rot is compartmentalization, where trees seal off damaged parts, blocking access to the rest of the tree. Another one is the inherent capability of trees to resist or slow down the spread of rot by the presence of toxic or fungistatic secondary metabolites.

1.2.3. Heartwood rot resistance

As mentioned, the resistance of HW against rot has been ascribed to the presence of secondary metabolites [39]. In addition to rot resistance, the secondary metabolites reduce shrinkage and swelling of wood at different moisture levels, which also contribute to the durability of wood in service in outdoor environments [40]. Several of these secondary metabolites exhibit strong antimicrobial activity that acts as a defense against rot and fungus [41–43]. Taiwania (*Taiwania cryptomerioides* Hayata) HW contains phenolics and terpenoids [8,9] which have antifungal [10], antitermite [6], antibacterial [11], antitumor [12,13], and anti-inflammatory [14] activities [44]. The HW of white-cedar (*Thuja occidentalis L.*) contains several sesquiterpene alcohols [45], and the HW of Chinese juniper (*Juniperus chinensis L.*) harbors many different diterpenes [46]. Previous reports show that the monoterpenoids nootkatin and carvacrol are linked to rot resistance in yellow cedar (*Chamaecyparis nootkatensis* D.Don 1824) [47]. The sesquiterpene alcohol cedrol provides resistance to termite decay in eastern red cedar (*Juniperus virginiana* L.) and western juniper (*Juniperus occidentalis* Hook.) [48,49]. HW rot resistance in European larch (*Larix decidua* Mill.), *P. sylverstris* [50] and *T. cryptomerioides* [51] is strongly correlated with high concentration of phenolics. Similar

correlations have also been found in angiosperm trees. Examples are, phenolic compounds such as stilbene in osage orange (*Maclura pomif*era Raf. Schneid.) [52], tannins and lignans in black locust (*Robinia pseudoacacia* L.) [53], flavonoids, naphthoquinones, and hydrolysable tannins in walnut (*Juglans nigr*a Linnaeus) [56], sesquiterpene alcohols such as catalponol in southern catalpa (*Catalpa bignonioides* Walter) [55]. In honey mesquite (*Prosopis glandulosa* Torr.) alkaloids are reported to associate with HW resistance [56], which is unusual as alkaloids are better known for their toxicity to the central nervous system of mammals [57].

1.2.4. Thuja plicata heartwood rot

As mentioned above, *T. plicata* trees can have long lifespans and the wood is highly valued for its natural durability, indicating considerable HW rot resistance. Paradoxically, substantial volumes of harvested logs are culled due to extensive rot (Figure 1.2) [2] and rot resistance is low relative to sympatric conifer species. A comparison of old growth trees in Southern Alaska showed that 100-year-old harvested *T. plicata* trees, the culled volume surpasses that of Western hemlock (*Tsuga heterophylla* Raf. Sarg.) and Sitka spruce (*Picea sitchensis* Bong. Carr.), and in 300-year-old trees, 50% of volumes are culled in *T. plicata*, whereas about 8% are culled in *T. heterophylla* and 5% are culled in *P. sitchensis* [58]. The same author showed that at age 100, 40% of *T. plicata* trees are afflicted by rot whereas 20% of *T. heterophylla* trees and 0% of *P. sitchensis* are afflicted by rot. Second-growth trees are similarly afflicted by rot [59]. As old-growth *T. plicata* -containing forests are rapidly disappearing in BC, the resistance in younger trees appear to vary between and within trees [59,60] and this variation is thought to be associated with levels of HW secondary metabolites. Therefore, the amount of secondary metabolites is used as a proxy of durability [61]. The highest amounts of these are found in the HW adjacent to the SW near the base of the tree and the lowest amount are found near the pith [62]. Research has found that secondary metabolite levels are indeed heritable traits [63]. There is evidence that HW rot resistance in growing trees of *T. plicata* correlates with high levels of terpenoid tropolone (e.g., β-thujaplicin) and that high levels of lignans correlate with rot resistance in wood in service [10,60]. The identity of other classes of secondary metabolites and their potential contribution to HW rot resistance in *T. plicata* is currently unknown.

Figure 1.2. T. plicata logs at a sorting station, Vancouver Island.

The amount of HW rot is so extensive that whole logs are culled. Blue bottle is approximately 35 cm high.

1.3. Resistance mechanisms

Plants are subjected to a variety of biotic and abiotic stresses throughout their long life. Biotic stresses range from bacteria, fungi, to herbivorous insects and mammals. Abiotic stresses include salinity, drought, UV radiation, heat, and cold [64,65]. Plants have evolved constitutive and inducible defenses against these stressors [64]. Constitutive biotic defences act as first barriers and include structures such as bark, thorns and trichomes and extra cellular secretion of lignin, cutin, and suberin substances that are hard to degrade, and secretion of terpenoids into glands and ducts. Induced defences include increased production of mentioned secreted substances, additional secondary metabolites, pathogen-degrading enzymes, and even cell suicide a.k.a. hypersensitive response, to remove substrate for pathogen spread [66–69]. Moreover, the constitutive and induced defences can be further classified into direct and indirect defences. Direct defences include morphological structures and secondary metabolites. On the other hand, indirect defences include plant volatiles, elicitors, phytohormones and natural enemies of herbivores [64,67,70,71]. Studies have shown many proteins and enzymes also play a significant role in defense. Such examples are α-amylase inhibitors, lectins, proteinase inhibitors, proline-rich proteins and pathogenesis-related (PR) proteins, i.e., hydrolytic enzymes such as glucanases, chitinases, and other hydrolases.

Others are systemic acquired resistance (SAR), programmed cell death known as hypersensitive response and compartmentalization that creates a chemical and anatomical boundary at the infected site. As such, one of the most thoroughly studied defense tools is secondary metabolites.

1.4. Secondary metabolites

Secondary metabolites are organic compounds produced by bacteria, fungi, and plants [72]. Unlike primary metabolites, secondary compounds are not involved in normal growth and development, instead play a role in defense as well as in ecological interactions [68]. These metabolites can be either constitutively present as inactive or compartmentalized forms or induced upon insect or microbe attack. The former are known as phytoanticipins and the latter as phytoalexins [64]. More than 300,000 different plant secondary metabolites have been characterized [73,74]. According to chemical structures, secondary metabolites are classified into three major groups i.e., terpenes, phenolics, and nitrogen-containing compounds [75]. Among them, terpenes are the largest and most diverse class of plant secondary compounds.

1.5. Other classes of secondary metabolites

In addition to terpenes, other classes of secondary metabolites include phenolic compounds (i.e., tannins, coumarins, flavonoids, chromones, xanthones, stilbenes, and lignans), nitrogen containing compounds (i.e., alkaloids, cyanogenic glycosides, glucosinolates, and nonprotein amino acids), saponins, lipids (i.e., fixed oils, waxes, essential oils, sterols, fat-soluble vitamins, and phospholipids) and carbohydrates [75]. Several studies have reported their role in defense as well as in other processes. Tannins (i.e., hydrolysable, and condensed tannins) have protein precipitating abilities, and have been used for decades to convert raw animal hides into leather. They can protect the heartwood from fungal and bacterial decay. Colorful flavonoids such as anthocyanins attract pollinators; flavones and flavonols protect cells from ultraviolet radiation; and isoflavones, which show antimicrobial activity against bacteria and fungi. Alkaloids are toxic to bacteria and fungi, can disrupt the protein activity after ingestion, and can alter signaling in the nervous system of various animals, including humans [76]. Since the focus of this thesis is the genetic basis of terpene defenses in *T. plicata*, I will discuss their biosynthesis in some detail below.

1.6. Terpenes

The name “terpene” is derived from the Latin word “turpentine”, referring to the sticky resin collected typically from pine trees and used for various purposes for thousands of years [77]. Terpenes make up the largest group of secondary metabolites (over 50,000 substances). From a chemical perspective, terpenes are also referred to as terpenoids (or isoprenoids). Terpenes are simple hydrocarbon structures, while terpenoids are modified terpenes with an oxygen moiety and additional structural rearrangements. However, these two terms are often used interchangeably [78]. All terpenes are derived from the 5-carbon isoprene units, isopentenyl diphosphate (IPP) and its isomer dimethylallyl diphosphate (DMAPP) [8]. Subsequent condensation of IPP and DMAPP by prenyltransferases (PTs) produces geranyl diphosphate (GDP, C_{10}), farnesyl diphosphate (FDP, C_{15}), and geranylgeranyl diphosphate (GGDP, C_{20}), which are considered universal terpenoid precursors. Additional terpenoid precursors such as neryl diphosphate (NDP, $C_{10,}$ cis isomer of GDP), as well as lavandulyl diphosphate (LDP) and chrysanthemyl diphosphate (CDP) which catalyzes irregular monoterpenoids have also been identified [79,80,81]. Specific synthases, collectively known as terpene synthases (TPSs) act on these universal precursors to produce a combinatorial diversity of terpenes [82,83]. Depending on the number of carbon atoms they possess, terpenes are classified as hemiterpenes (C_5), monoterpenes (C_{10}), homoterpenes ($C_{11,16}$), sesquiterpenes (C_{15}), diterpenes (C_{20}), sesterterpenes (C_{25}), triterpenes (C_{30}), sesquarterpenes (C_{35}), tetraterpenes (C_{40}), and polyterpenes ($C_{>40}$) [78]. Some terpenes deviates from the general 5n carbon backbone, such as homoterpenes (C_{11} and C_{16}), sterols and steroids ($C_{27\text{-}29}$), referred to as irregular terpenoids [84].

1.6.1. Functions in plants

Based on general toxicity, terpenoids are thought to play important roles in biotic stress defenses in the plant kingdom [85]. Terpenoids are well known for the defense against herbivores and pathogens in gymnosperms and angiosperms species [86,87]. In addition to toxicity, they can also act as allelopathic agents, feeding deterrents, and pollinator attractants. Such example includes monoterpene pinene, which is toxic to the worldwide damage caused by bark beetles in conifers. Other monoterpenoids such as linalool and 1,8-cineole serve as attractants for pollinators, while camphor acts as a feeding deterrent to herbivores and can also act as an allelopathic agent by preventing

seed germination of other species [88,89]. Numerous sesquiterpenoids act as phytoalexins and antifeedants that discourage opportunistic herbivory. Sesquiterpenoid capsidiol from tobacco (*Nicotiana tabacum* L.), lubimin and rishitin from potato (*Solanum tuberosum* L.), and gossypol from cotton (*Gossypium arboretum* L.) all act as phytoalexins [90–92], while polygodial from water pepper (*Persicaria hydropiper* L. Delabre) act as feeding deterrent of herbivorous insects [93]. Diterpene hydrocarbon casbene synthesized by casbene synthase from castor bean (*Ricinus communis* L.) acts as a phytoalexin [94]. Diterpenes confer resistance against pests and pathogens in particularly well-studied model plant species, including rice (*Oryza sativa* L.), maize (*Zea mays* L.) [95,96], *N. tabacum*, [97] and Arabidopsis (*Arabidopsis thaliana* L. Heynh). [98]. Mutants defective in the biosynthesis of these compounds are sensitive to infections, providing more direct evidence for a role in biotic defenses than correlations between levels of compounds and disease resistance in other species. Terpenoids serve important physiological and ecological functions in plants, by acting, for example, as growth regulators, photosynthetic pigments, electron transporters or by facilitating plant-to-plant communication, or by providing membrane stability [99]. To name a few, sesquiterpene abscisic acid and diterpene gibberellins are growth regulators which regulate seed dormancy, seed germination, stomatal closure, stem elongation and flower induction [100]. Tetraterpenoids, chlorophyll and carotenoids are photosynthetic pigments or chromophores, crucial for photosynthesis [101]. For example, the volatile emissions of monoterpenes α-pinene, β-pinene, and camphene from sender plants induced defense in neighboring plants [102]. Polyterpenoids, plastoquinones, and ubiquinones, function as electron transporters [103]. Triterpenoid sterols i.e., phytosterols promote membrane stability [104]. Hemiterpenoid isoprene prevents cellular damage at high temperatures, perhaps by reacting with free radicals, stabilizing membrane components [105]. Many terpenoids provide aroma, flavors and colors to plants. To name a few, monoterpenoid eucalyptol accounts for the mint-like smell of eucalyptus (*Eucalyptus globulus* Labill.) [106] while the sweet and fruity aroma of mango (*Mangifera indica* L.) fruit depends on monoterpenes like terpinolene, 3-carene, caryophyllene, α-pinene and β-myrcene [107]. Valencene is a sesquiterpene, an aroma component of citrus fruit and its derivative, nootkatone, is one of the main components of the aroma and flavor of grapefruit. The flavors of cinnamon (*Cinnamomum verum* J.Presl), cloves (*Syzygium aromaticum* L. Merr. & L.M.Perry), and spices like oregano (*Origanum vulgare* Linnaeus*)*, basil (*Ocimum basilicum* L.) and rosemary (*Salvia*

rosmarinus Spenn.) come from sesquiterpene caryophyllene [108]. Tetraterpenoids β-carotene and lycopene are natural carotenoids and impart orange and red colors to fruits and vegetables [109]. Certain terpenoids may also be involved in plants' protection against abiotic stresses, such as exposure to high temperatures, and reactive oxygen species-derived oxidative damage [110].

1.6.2. Human uses

Many terpenoids have pharmaceutical properties and are used to treat human diseases [82]. They have played a long role in traditional medicine for various purposes and are also used in Western medicine. Well-known examples are monoterpenoid camphor and cineole (also called eucalyptol), which are used as analgesic drugs to relieve pain caused by burns, strains, and other inflammations. Several monoterpenes, such as pinene, 3-carene, as well as monoterpene alcohol geraniol and thymol possess numerous beneficial medicinal properties, including antioxidant, anti-inflammatory, antimicrobial, and anticancer activities. Monoterpenoid tropolone β-thujaplicin (also called hinokitiol) found in Cupresssaceae heartwood has antifungal, antibacterial, antiviral, antiinflammatory and anticancer activity. The monoterpene alcohol menthol is used medicinally in ointments, cough drops, and nasal inhalers [111]. It is also used as a flavouring in foods, cosmetics, and perfumes. The sesquiterpene lactone artemisinin is used as an effective anti-malarial drug obtained from the traditional Chinese medicinal plant sweet wormwood (*Artemisia annua* L.), a plant used in medicine since 200 B.C [112]. Terpenoids have been used in chemotherapy of various cancers, with taxol (a.k.a. paclitaxel) obtained from the bark of Pacific yew (*Taxus brevifolia* Nutt.), a highly modified diterpenoid, being the to date most commercially successful compound used in cancer treatment [113]. Another diterpenoid, forskolin, obtained from the root of an aromatic herb (*Coleus barbatus* Andrews Benth. ex G.Don), is used in eye drops to treat glaucoma [114]. Diterpenoids such as tanshinones, which are abietane-type norditerpenoid quinones obtained from the Chinese medicinal herb red sage (*Salvia miltiorrhiza* Bunge) have various biological activities, including anticancer, anti-inflammatory, antimicrobial, antivirus, antioxidant, and neuroprotective activities [115]. Triterpenoid steroids and saponin are used as anti-inflammatory and anticancer drugs [116]. Tetraterpenoids such as carotenoids, namely β-carotene is used as antioxidants, protecting cells from cancer and other diseases [117]. Terpenoids are also used in the chemical, agriculture, and food industries. For example, monoterpenoid tropolone β-

thujaplicin is widely used in skin and oral care products such as soaps, lotions, hair tonics and toothpastes. It is also used as food and wood preservatives [118]. Monoterpenoid camphor is one of the most commercially important aroma chemicals with an annual market value of US$80–100 million [119]. Monoterpene geraniol and linalool from roses and orchids are extensively used in the fragrance and cosmetics industries because of their pleasant rose-like scent [120]. A monoterpene ester called pyrethroid is used as an active ingredient in commercial biopesticides [121]. Tropical *S. album* produces one of the world's most highly prized fragrances, which are sesquiterpenes α-santalol and β-santalol [122]. Triterpenoid saponins are biosurfactants and are used as soap, hair cleanser and detergent [123]. Monoterpenes like limonene, linalool, 1,8-cineole are used for the aroma of lemon or lime beverages [124]. Tetraterpenoids such as carotenoids are used as natural colorants in food, drinks and cosmetics and are a natural source of vitamins [125]. Terpenoids have recently been exploited in the development of biofuel products. Monoterpenes limonene and β-phellandrene, the sesquiterpenes farnesene and bisabolene and the diterpenes phytene and cambrene are used as replacements for transportation fuels [126].

1.6.3. Terpene Backbone Biosynthesis

The isoprene backbone of terpenes is synthesized by two compartmentalized pathways: the methylerythritol 4-phosphate (MEP) pathway present in plastids and the mevalonate (MVA) pathway present in the cytosol (Figure 1.3) [127,128].

In cytosol, the MVA pathway consists of six steps and starts with the condensation of two molecules of acetyl-CoA to acetoacetyl-CoA (ACAC-CoA) catalyzed by acetoacetyl-CoA thiolase (AACT) [129]. ACAC-CoA is combined with a third molecule of acetyl-CoA to form 3-hydroxy-3-methylglutaryl-CoA (HMG-CoA) which is catalyzed by HMG-CoA synthase (HMGS) [130]. In the following rate-limiting step, HMG-CoA reductase (HMGR) catalyzes the conversion of HMG-CoA to mevalonic acid (MVA) [131]. MVA is finally converted into IPP via three enzymatic steps: two ATP-dependent phosphorylation steps, catalyzed by mevalonate kinase (MK) and phosphomevalonate kinase (PMK), and an ATP-driven decarboxylative elimination catalyzed by mevalonate-5-diphosphate decarboxylase (MDD). IPP is interconverted into DMAPP by isopentenyl diphosphate isomerase (IDI) [84]. Condensation of IPP and DMAPP reactions facilitated by prenyl transferases (PTs) generates farnesyl diphosphate (FDP), which is the

precursor substrate used by terpene synthases (TPSs) to produce sesquiterpenes, triterpenes, polyterpenes, phytosterols and brassinosteroids [99].

In plastid, the MEP pathway consists of seven enzymatic steps. In the first reaction, pyruvate reacts with thiamine diphosphate (TPP) to yield hydroxyethyl-TPP, which condenses with glyceraldehyde-3-phosphate (GAP) to form 1-deoxy-D-xylulose 5-phosphate (DXP) catalyzed by DXP synthase (DXS) [127]. The enzyme 1-deoxy-D-xylulose 5-phosphate reductoisomerase (DXR) catalyzes the second step of the MEP pathway, in which DXP is converted into 2-C-methyl-Derythritol 4-phosphate (MEP) [132,133]. MEP is further converted into 4-diphosphocytidyl-2- C-methyl-D-erythritol (CDP-ME) by the enzyme 4-diphosphocytidyl-2-C-methyl-Derythritol synthase (MCT). Phosphorylation of CDP-ME by the enzyme 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase (CMK) then leads to the formation of 4-diphosphocytidyl-2-C-methyl-D-erythritol 2-phosphate (CDPME2P), which is subsequently cyclized by 2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase (MDS) into 2-C-methyl-D-erythritol 2,4-cyclodiphosphate (MEcDP). In the last two steps of the MEP pathway, the enzyme 4-hydroxy-3-methylbut-2-enyl diphosphate synthase (HDS) first converts MEcDP into 4-hydroxy-3-methylbut-2-enyl diphosphate (HMBPP) [82]. In a final branching step, HMBPP is converted by 4-hydroxy-3-methylbut-2-enyl diphosphate reductase (HDR) to IPP. Like in the MVA pathway, IPP is interconverted into DMAPP by isopentenyl diphosphate isomerase (IDI) [134]. Condensation of IPP and DMAPP reactions facilitated by prenyl transferases (PT), generates geranyl diphosphate (GDP), and geranylgeranyl diphosphate (GGDP), which are the precursor substrates used by terpene synthases (TPSs) to produce hemi-, mono-, di-, sester-, tetra-terpenes, phytohormones such as cytokinins and gibberellins, chlorophyll, tocopherols, and plastoquinones [135]. Besides TPSs, enzymes such as phytoene synthase (PHY) in the MEP pathway use GGDP as a precursor to form tetraterpenes such as carotenoids, whereas enzymes i.e., squalene synthase (SQS), squalene epoxidase (SQE) and triterpene synthase (TTS) in the MVA pathway, use FDP as a precursor to form triterpenes [136]. Though MVA and MEP pathways operate independently in different subcellular compartments, metabolic cross-talk between the two pathways has been reported [84].

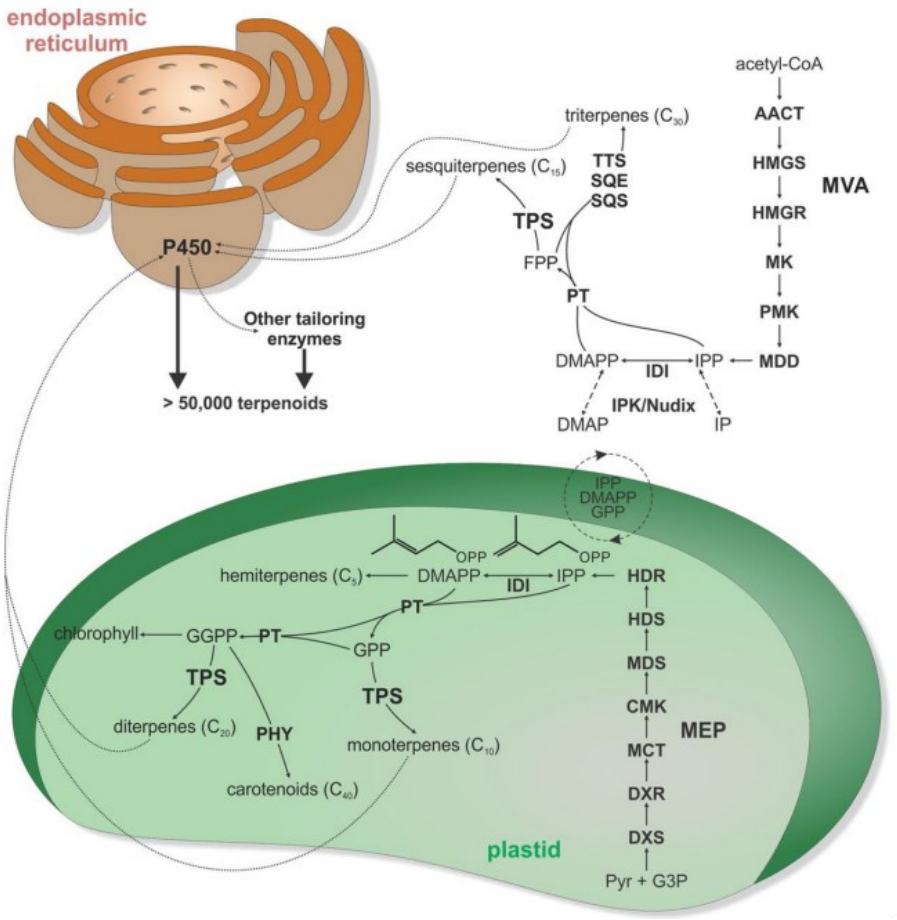


Figure 1.3. Terpenoid biosynthesis pathway.

Schematic overview of terpenoid biosynthetic pathways. All terpenoids are derived from two isomeric 5-carbon precursors, isopentenyl diphosphate (IPP), and dimethylallyl diphosphate (DMAPP). In turn, IPP and DMAPP are formed via two pathways, the cytosolic mevalonate (MVA) pathway originating from acetyl-CoA and the pyruvate and glyceraldehyde-3-phosphate (G3P)–derived 2-C-methyl-D-erythritol-4-phosphate (MEP) pathway located in the plastids. Condensation of IPP and DMAPP yields geranyl diphosphate (GPP) as the precursor to monoterpenoids (C_{10}), fusing GPP with an additional IPP results in the sesquiterpenoid (C_{15}) precursor farnesyl diphosphate (FPP), and fusing FPP with IPP generates geranylgeranyl diphosphate (GGPP) a route to diterpenoids (C_{20}). Furthermore, condensation of two FPP or two GGPP molecules forms the central substrates of triterpenoid (C_{30}) and carotenoids (C_{40}), respectively. Terpene synthases (TPS) are key gatekeepers in the biosynthesis of C_{10}–C_{20} terpenoids, catalyzing the committed scaffold-forming conversion of the respective prenyl diphosphate substrates into a range of hydrocarbon or oxygenated structures. These TPS products can then undergo various oxygenations through the activity of cytochrome P450 monooxygenases (P450), followed by further possible functional decorations, ultimately giving rise to more than 80,000 distinct natural products. AACT, acetoacetyl-CoA thiolase; CMK, 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase; DXR, 1-deoxy-D-xylulose 5-phosphate reductase; DXS, 1-deoxy-D-xylulose 5-phosphate synthase; HDR, (E)-4-hydroxy-3-methyl-but-2-enyl diphosphate reductase; HDS, (E)-4-hydroxy-3- methyl-but-2-

enyl diphosphate synthase; HMGR, 3-hydroxy-3-methylglutaryl-CoA reductase; HMGS, 3-hydroxy-3-methylglutaryl-CoA synthase; IDI, isopentenyl diphosphate isomerase; MCT, MEP cytidyltransferase; MDD, mevalonate-5-diphosphate decarboxylase; MDS, 2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase; MK, mevalonate kinase; P450, cytochrome P450-dependent monooxygenase; PHY, phytoene synthase; PMK, phosphomevalonate kinase; PT, prenyl transferase; SCS, squalene synthase; SQE, squalene epoxidase; TPS, terpene synthase; TTS, triterpene synthase. Unaltered figure reprinted from Karunanithi, P.S.; Zerbe, P. Terpene Synthases as Metabolic Gatekeepers in the Evolution of Plant Terpenoid Chemical Diversity. Front. Plant Sci. 2019, 10, 1–23, under a Creative Commons 4.0 agreement.

1.6.4. Cell specialization in terpenoid biosynthesis

Terpenoids are produced and stored in highly specialized structures [137]. In angiosperms species, they are stored in glandular trichomes, glandular epidermis and secretory cavities. In gymnosperms they are stored in tube-like resin ducts and sac-like blisters, as monoterpene olefins (turpentine), sesquiterpenes and diterpene resin acids (rosin), present in seeds, needles, bark, phloem, and xylem [138]. Since terpenoids have low solubility in water and are toxic to plant cells, they are typically synthesized in specialized cells that secrete them into extra-cellular specialized structures or cell wall spaces to prevent autotoxicity and to separate secondary from primary metabolic processes. Research has shown that these specialized cells are primarily ray cells, but also specialized epithelial cells in Pinaceae species and axial xylary parenchyma cells in Cupressaceae species [138–142]. These are living cells present in SW as the developmentally regulated SW to HW transition occurs [12,13,140,143]. Stored terpenoids appear to exert their toxic effect by diffusing into cell wall spaces, by volatilization to deter herbivores, and by glands breaking and releasing their content upon herbivory.

1.6.5. Terpene synthases

The chemical diversity of terpenes is driven by the stereo-specific carbocation cyclization, rearrangement, and elimination reactions that transform the precursors into distinct terpenes. This indicates that a single terpene synthase (TPS) enzyme can give rise to multiple products due to the stochasticity of bond rearrangement following the generation of the unusual carbocation intermediates. TPS enzymes can produce a wide range of mono-, sesqui- and diterpenes from their respective geranyl diphosphate (GDP), farnesyl diphosphate (FDP) and geranyl geranyl diphosphate (GGDP) substrates [144,145]. The TPS gene families are larger in conifers than in other organisms. Based

on transcriptome data of several spruce (*Picea*) species, spruce genomes contain on average at least 69 unique and transcriptionally active TPS genes. In contrast, the Arabidopsis and poplar genomes contain about 32 and the rice genome contains about 31 TPS genes [146–149]. Phylogenetic analyses suggest that these genes evolved from a combination of gene duplication, neo-functionalization and/or sub-functionalization events [150,151]. The *T. plicata* genome was recently sequenced, resulting in an estimated genome size of 9.8GB and 39,659 gene models. Based on Gene Ontology analysis about 140 transcripts were annotated as encoding TPSs [152]. The exact structural features that determine the diverse product profiles of TPS enzymes are not known [134,146,150]. A diverse array of monoterpene synthases (monoTPSs), sesquiterpene synthases (sesquiTPSs) and diterpene synthases (diTPSs) have been cloned and functionally characterized in angiosperms and gymnosperm species [150,153–155]. All monoTPS enzymes use a common reaction mechanism in which ionization of GDP forms linalyl diphosphate (LDP). Ionization of an enzyme-bound LDP intermediate promotes cyclization to a six-membered ring carbocation called α-terpinyl cation which undergoes additional electrophilic cyclizations, hydride shifts and other rearrangements. The reaction is terminated by either deprotonation of the carbocation intermediate or capture by a nucleophile, e.g., water. As mentioned above, monoTPSs can produce more than one product. For instance, pinene synthase from Pineaceae can produce both α & β-pinene. MonoTPSs can produce acyclic (i.e., myrcene), monocyclic (i.e., limonene) and bicyclic (i.e., α, β-pinene) products from GDP. The same enzymes can also transform GDP to oxygenated products such as camphor, 1,8-cineol as well as alcohols (i.e., linalool, geraniol, thymol) and esters (i.e., pyrethrin). The reaction mechanism for sesquiterpene synthases (sesquiTPSs) closely resembles that of monoTPSs. The additional C_5 unit permits FDP to generate a greater number of skeletal structures of sesquiterpenes compared to GDP. Diterpene synthases (diTPSs) catalyze two different types of reactions. The first reaction type resembles to mono- and sesquiTPSs. Products generated from such reactions are casabene and taxadiene are catalyzed by casabene and taxadiene synthases. In the second type of reaction, cyclization of GGDP by diTPSs forms labdadienyl diphosphate, also known as copalyl diphosphate (CDP), followed by a second cyclization carried out by the same enzyme to make characteristic diterpenes.

These types of diTPSs are known as bifunctional [156] and an example is abietadiene catalyzed by abietadiene synthase from Grand fir (*Abies grandis* Douglas ex D. Don Lindley). Other diTPSs have a complementary activity, converting CDP into specific diterpenes [157,158]. DiTPSs can make acyclic, bicyclic (labdanes, clerodanes), tricyclic (pimaranes, abietanes, cassanes, rosanes), tetracyclic or macrocyclic types of compounds. Oxidation of diterpenes forms resin acids such as abietic acid, important for wound sealing in conifers which are subsequently transformed into amber. The substrate for triterpene is squalene produced by condensation of two molecules of FDP catalyzed by triterpene synthase, i.e., squalene synthase. Oxidation of squalene forms oxidosqualene. The substrate for tetraterpene is phytone produced by condensation of two molecules of GGDP catalyzed by tetraterpene synthase, i.e., phytone synthase that produces β-carotene and lycopene.

Mono-, sesqui- and diterpene synthases are characterized by conserved functional motifs (Figure 1.4) [135]. MonoTPSs have an N-terminal transit peptide for importing in plastids, whereas sesquiTPSs lack N-terminal transit peptide sequences, consistent with cytosolic localization of plant sesquiTPSs. MonoTPSs and some sesquiTPSs contain conserved N terminal RRX_8W (where “X” can be any amino acid) motif and the highly conserved C-terminal aspartate-rich DDXXD motif. The RRX_8W motif is required for the initial isomerization of GDP into linalyl diphosphate and required for the subsequent cyclization in cyclic monoterpenes [159–161]. Studies have shown that deletion of the RRX_8W motif from monoterpene synthase can affect the ability of the enzyme to utilize GDP as a substrate [161]. The DDXXD motif is crucial for divalent cation (typically Mg^{2+} or Mn^{2+}) assisted substrate binding and ionization [162–164]. The mutation of the DDXXD motif leads to decreased catalytic activity and formation of abnormal product [165]. Depending on domain structure, active sites, and signature motifs, there are three major classes of diTPSs: monofunctional class I and class II diTPSs and bifunctional class I/II diTPSs [166,167]. Class II diTPSs contain a conserved DXDD motif critical for protonation-initiated cyclization of geranylgeranyl diphosphate (GGDP) into the bicyclic prenyl diphosphates *ent*-copalyl diphosphate (*ent*-CDP), normal-CDP, or *syn*-CDP [168,169]. In class I diTPSs, the conserved DDXXD motif in the C-terminal domain is required for the diphosphate ionization-initiated cyclization and rearrangement of CDP to diterpene hydrocarbons [170].

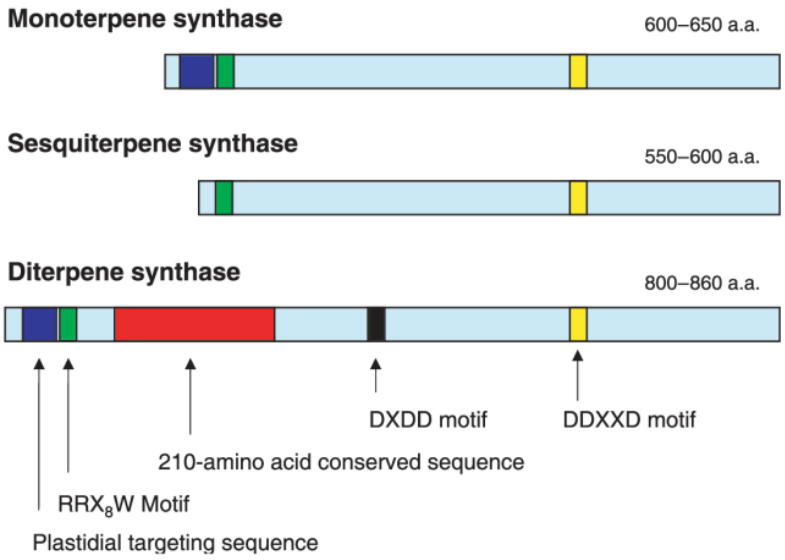


Figure 1.4. Layout of mono-, sesqui-, and diterpene synthase protein sequences in conifer.

Schematic illustration of the motif profiles of mono-, sesqui- and diterpene synthases. Reprinted from Genes, enzymes, and chemicals of terpenoid diversity in the constitutive and induced defence of conifers against insects and pathogens, by C.I. Keeling and J. Bohlmann, 2006, New Phytologist, 170(4), p. 660. Reprinted with permission from the publisher.

Bifunctional class I/II diTPSs combine both reactions in a single protein where CDP intermediate freely diffuses from class II to the class I active site [171]. The function of diTPS enzymes can be altered by the steric and electrostatic control of the intermediate carbocations formed through substrate protonation (class II) or ionization (class I) [172,173]. A single amino acid substitution can lead to the redirection of diTPS product specificity, highlighting the functional plasticity of diTPSs and their capacity to evolve new functions [151,174–176].

Several genes encoding synthases for mono-, sesqui-, and diterpenes have been identified in conifer species based on their potential roles in insect resistance [177–181]. To date, only one *T. plicata* monoTPS gene has been cloned. It is expressed in the epithelium of foliar glands and produces sabinene, the predicted precursor for the well-known oxygenated monoterpenoid α-thujone [182].

1.6.6. Modification of terpenes

Terpenes can undergo various enzymatic modifications, such as oxygenation, hydroxylation, dehydrogenation, acylation, epoxidation, followed by further possible

functional decoration and structural rearrangements, giving rise to more than 200,000 distinct natural products. These modifications are performed by oxidases such as cytochrome P450 oxygenases and Fe(II)/2-oxoglutarate dependent dioxygenases, as well as dehydrogenases, isomerases, reductases, methyltransferases, acyltransferases, and glycosyltransferases or in some cases by terpene synthases. Chemical analyses show, however, that in various organisms, terpenes are found as terpene-like or terpenoids, primarily with various forms of oxygenated side groups. The terpenoids occur as alcohols, esters, aldehydes, acids, and ketones, leading to a considerable diversity of terpenoids, and presumably attached diversity of chemical and biological activities. Over the last few decades, over 80,000 chemical structures of terpenoids have been discovered [183]. To date, different structural types of >1000 monoterpenes, >7000 sesquiterpenes and >3000 diterpenes are defined [184]. In peppermint (*Mentha x piperita L.*), transformation of limonene is carried out by a microsomal cytochrome P450 limonene 3-hydroxylase to produce oxygenated alcohol, isopiperitenol, whereas a reductase converts a ketone, menthone, to an alcohol, menthol, a compound responsible for the characteristic flavor and cooling sensation of peppermint [185]. Microsomal preparations from fruits of caraway (*Carum carvi* L.) catalyze the hydroxylation of limonene to monoterpenoid carveol, the key intermediate in the biosynthesis of carvone [186]. Diterpenes are typically found as carboxylic acids in conifers, better known as diterpene resin acids [96]. Biochemical studies provide evidence for modification of diterpenes first by hydroxylation into alcohols followed by oxidation into aldehydes and acids [187]. In diterpene biosynthesis, cytochrome P450 enzymes are involved in both primary metabolism to make gibberellin and also in secondary metabolism to make diterpene resin acids [96]. Several conifer cytochrome P450-encoding genes have been cloned and shown to carry out diterpene alcohol oxidations into diterpene carboxylic acids [178,184,187–190]. *S. album* cytochrome P450s of the CYP76F sub-family were recently shown to hydroxylate santalenes and bergamotene to santalols and bergamotol [191]. *T. plicata* produces high amounts of oxygenated thujone monoterpenoids in foliage, which are associated with resistance against herbivore feeding, particularly ungulate browsing. Thujones are produced from sabinene by sabinol and sabinone enzymes. Two cytochrome P450 enzymes called sabinene hydroxylases catalyze the hydroxylation of sabinene to sabinol, the second step in the thujone pathway [192]. The other functionally characterized terpene-modifying P450 genes of gymnosperms are involved in taxol biosynthesis in *T. brevifolia*

[193], diterpene resin acid biosynthesis in loblolly pine (*Pinus taeda* Carl Linnaeus, 1753) [184] and *P. sitchensis* [178], and sesquiterpenoid nootkatone biosynthesis in *C. nootkatensis* [194]. In Mexican cypress (*Cupressus lusitanica* Mill.), oxidation of terpinolene is carried out by cytochrome P450 monooxygenases, which generate various oxygenated monoterpenoid precursors for tropolone β-thujaplicin biosynthesis [195].

1.6.7. Regulation of terpenoid biosynthesis

The biosynthesis of specialized terpenoids is regulated at various levels. Studies have shown that transcription factors (TFs), such as WRKY, myeloblastosis (MYB), Basic helix–loop–helix (bHLH), Ethylene Responsive Factor (ERF), Jasmonate Responsive ERF (JRE), Basic Leucine Zipper (bZIP) and more can regulate the biosynthesis of terpenoids [84]. This regulation is mainly controlled at the level of transcription through binding of TFs at specific promoter sequences of target genes [196]. For example, WRKY TFs regulate the biosynthesis of many terpenoids such as monoterpenoids in tomato (*Solanum lycopersicum* L.) [197], sesquiterpenoids artemisinin in *A. annua* [198], gossypol in *G. arboretum* [199] and capsidiol in tobacco (*Nicotiana benthamiana* Domin) [200], diterpenoids momilactone A in *O. sativa* [201] and taxol in Chinese yew (*Taxus chinensis* Rehder & E.H.Wilson Rehder) [202], and triterpenoids ginsenosides in ginseng (*Panax quinquefolius* L.) [203]. Besides TFs, the role of microRNAs in controlling specialized terpenoids biosynthesis has recently reported in *A. thaliana* [204]. Research has also shown that pathway intermediates and downstream metabolites can regulate the core terpenoid biosynthetic steps at both transcriptional and posttranslational levels. Since individual terpene synthases compete for the same substrate, the level of expression of their genes also influences the levels of corresponding products. Similarly, the expression level of genes required for modification of terpenes into specific terpenoids influence the amounts of corresponding terpenoid that is produced [192]. For example, the expression levels of 1-deoxy-D-xylulose 5-phosphate synthase (DXS), catalyzing the first step in the biosynthesis of isoprenes, can be rate limiting in the biosynthesis of terpenes in some species and processes [125,205]. Consequently, DXS mutants such as those of the single functional *A. thaliana* class-I type DXS gene (DXS1) exhibit albino phenotypes [32–34]. In *A. thaliana*, overexpression of the MVA pathway intermediate 3-hydroxy-3-methylglutaryl-CoA synthase (HMGS) not only elevated the expression of downstream gene 3-hydroxy-

3-methylglutaryl-CoA reductase (HMGR), but further down sterol biosynthesis is also upregulated [206]. A suitable supply of the building blocks for terpenoid biosynthesis, isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP) is extremely crucial for efficient terpenoid biosynthesis [132]. The conversion of IPP and DMAPP into their respective monophosphate forms IP and DMAP catalyzed by isopentenyl phosphate kinase (IPK) and Nudix hydrolase enzymes can also regulate terpenoid biosynthesis, possibly by modulating the ratios of their diphosphate forms [207]. As discussed earlier, terpenoid biosynthesis is regulated by herbivore feeding, pathogen attack, and wounding as well as by downstream hormones such as methyl jasmonate (MeJA) and ethylene [142]. For example, MeJA treatment mimics bark beetle attack, triggering the formation of traumatic resin ducts (TRD), and increasing the amount of terpenoids in bark and wood of *P. abies* [208]. MeJA also induces the accumulation of monoterpenoid β-thujaplicin in *C. lusitanica* cultures [209].

1.7. Thuja plicata heartwood secondary metabolites

Secondary metabolites extracted from *T. plicata* HW have been studied for some time. *T. plicata* heartwood synthesizes an array of secondary metabolites in response to biotic and abiotic stress. The most abundant secondary metabolites are monoterpenoid tropolones, and fungistatic lignans [61,145,210,211]. To date, at least eight tropolone and eleven lignan compounds have been identified and characterized in *T. plicata.* The identified tropolones are thujaplicin (in α, β and γ isoforms), β-thujaplicinol, β-dolabrin, thujic acid, methyl thujate and trace amounts of nezukone, and thujin [10,61,145]. Among these, the most abundant are β and γ-thujaplicins [61]. In *T. plicata*, a high concentration of thujaplicins are found in the outer heartwood, which decreased closer to the pith [212]. Thujaplicins are strong metal chelators [213–215] preventing iron-dependent fenton reactions during degradation of lignocellulosic cell walls [211,216]. The ketone structure in its molecule facilitates the formation of metal chelates in the presence of various metal ions. In addition to tropolones, two monoterpenes i.e., carvacrol and carvacrol methyl ether have also been identified in *T. plicata* heartwood extracts [217].

The second main group of *T. plicata* heartwood secondary metabolites are lignans such as plicatic acid, plicatin, thujaplicatin and thujaplicatin methyl ether, and seven other compounds at low quantities [10,61,145]. Plicatic acid accounts for 40-50%

of total secondary metabolites. It is a strong acid because of its highly hydroxylated side chain and is responsible for the red-brown color of *T. plicata* heartwood [218]. Plicatin, thujaplicatin methyl ether, thujaplicatin make up 30-40% of the total lignan content. Lignans are abundant in wood in service, tropolones are degraded over time and durability correlates primarily with lignan content [211]. Lignans have fungistatic activity [219]. Plicatic acid is a radical scavenger [220]. The identity of other terpenes in *T. plicata* heartwood is unknown.

1.7.1. Biosynthesis of β-thujaplicin

The tropolone β-thujaplicin is a natural monoterpenoid found in the wood of trees in the family of Cupressaceae [213,221]. It is characterized by a ring expansion from six to seven carbons, generating a non-benzene aromatic ring. This compound was first discovered in the heartwood of *T. plicata* [222–224] and exhibits strong antifungal and antitumor activities [118,224,225]. More than 80 years have elapsed since it was discovered and there is still a lack of information on its biosynthesis pathway. Some information has been gleaned by using *C. lusitanica* cultured cells as a model. The production of β-thujaplicin can be induced in these cell cultures by addition of fungal extracts or the hormone MeJA [209,226,227]. After elicitation, six non-volatiles oxidized monoterpenes such as 4- terpineol, α-terpineol, 4-hydroxyphellandric acid methyl ester, 1,6-epoxy-4(8)-p-menthen-2-ol, p-ment-4(8)-en-1, 2-diol and β-thujaplicin and 10 volatiles hydrocarbon monoterpenes such as α-pinene, sabinene, β-pinene, myrcene, p-cymene, α-terpinene, limonene, βocimene, γ-terpinene and terpinolene have been identified from the ether extracts of cells and medium [227]. Radioisotope feeding experiments in these cell cultures have shown that geraniol and glucose were efficiently incorporated into β-thujaplicin via the MEP pathway, suggesting that the MEP pathway acts as a main pathway for β-thujaplicin biosynthesis, although the MVA pathway also contributes a minor portion of β-thujaplicin biosynthesis [222,228]. A terpene synthase assay revealed that terpinolene is the first olefin produced from GDP. Subsequently, feeding cell cultures with deuterium-labelled terpinolene further pinpoints terpinolene as an intermediate metabolite in β-thujaplicin production [223]. Terpinolene was the principal monoterpene produced, and thereafter rapidly metabolized into the ring-expanded tropolone β-thujaplicin [227]. Feedings of suspension culture with 3-carene and terpinyl acetate significantly improved the production of β-thujaplicin, indicating that these two compounds could act as alternative substrates in biosynthesis of β-thujaplicin

[229]. More investigations on this cell culture have shown that microsomal fractions from cells oxidized terpinolene into the hydroxylated compound 5-isopropylidene-2-nethylcyclohex-2-enol (IME) as a sole reaction product. A second microsomal reaction revealed that IME was selectively oxidized into the epoxidized compound 1,6- epoxy-4(8)-p-menthen-2-ol (EMO) to generate β-thujaplicin [195,230]. Experiments on kinetics and with specific inhibitors confirmed that these reactions were caused by cytochrome P450 monooxygenases, respectively [195]. The enzymatic function required for the ring expansion of EMO into β-thujaplicin is still unknown and the intermediate between EMO and β-thujaplicin biosynthesis is also unknown (Figure 1.5).

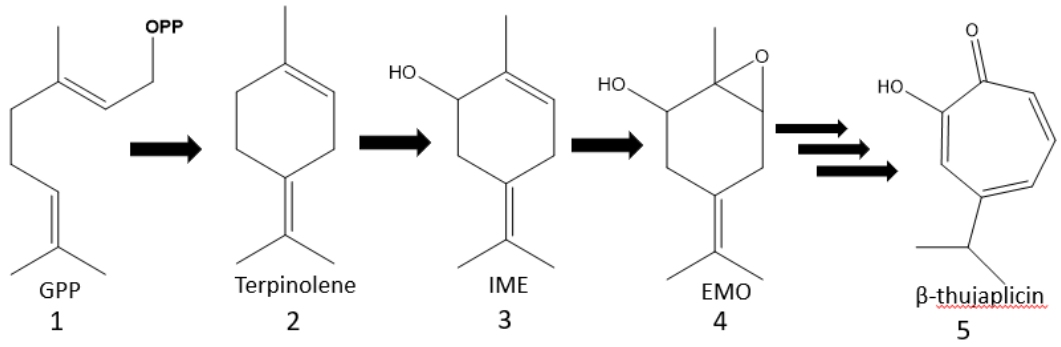


Figure 1.5. Predicted β-thujaplicin biosynthesis pathway.

1) GPP (geranyl diphosphate) a.k.a. GDP 2) Terpinolene 3) IME (5-isopropylidene-2-nethylcyclohex-2-enol) 4) EMO (1,6- epoxy-4(8)-p-menthen-2-ol) 5) β-thujaplicin. Adapted from Fujita, K.; Bunyu, Y.; Kuroda, K.; Ashitani, T.; Shigeto, J.; Tsutsumi, Y. A Novel Synthetic Pathway for Tropolone Ring Formation via the Olefin Monoterpene Intermediate Terpinolene in Cultured Cupressus Lusitanica Cells. J. Plant Physiol. 2014, 171, 610–614.

Tropolone biosynthesis pathways have been elucidated only in two fungal [231] and one bacterial [232] species. The pathways in plants have not been uncovered yet except for the alkaloid compound colchicine present in the seed of the flower autumn crocus (*Colchicum autumnale* L.). However, the origin and the pattern of ring formation is different from those of β-thujaplicin [233,234]. There is evidence from work with two related fungal species (*Penicillium stipitatum*, *Penicillium puberulum*) that Fe(II)-dependent dioxygenase enzymes catalyze the oxidative ring expansion of cyclohexadienone from six to seven carbons to produce fungal tropolones e.g., stipitatic acid and puberulic acid [231]. In *E. coli*, an epoxide hydrolase enzyme converts spiroepoxide into 6-hydroxy-6-hydroxymethylcyclohexa-2,4-dienone through an oxidative dearomatization process. In the presence of Fe(II)-dependent dioxygenase, 2,4-dienone

undergoes pinacol-type rearrangements resulting in ring extension reaction to form 2-tropolone [232]. In bacteria, phenylacetic acid is the precursor of the seven-member ring tropolone, which itself is derived either directly, or indirectly via phenylalanine, from the shikimate pathway [232], while fungi use a polyketide route [231] and in plants, originates from alkaloids and terpenes is known [223,234].

1.7.2. Fe(II)/2-oxoglutarate dependent dioxygenases

Oxygenase enzymes catalyze the formation of oxygenated compounds through C−H activation reactions. Oxygenase enzymes are divided into monooxygenases and dioxygenases. The monooxygenase category includes the cytochrome P450s and flavin-dependent oxygenases, where one oxygen atom is incorporated into the product and the other is being reduced to water by the cofactor NADPH. Whereas the dioxygenase category includes non-heme iron (II)/2oxoglutarate dependent dioxygenases (2OGDs), which incorporates both oxygen atoms into the product [235]. In plants, apart from cytochrome P450s, the 2OGD superfamily constitutes the second-largest family of enzymes [73,236]. They are non-heme iron-containing proteins, localized as a soluble protein in the cytosol. The catalytic core of these enzymes contains the conserved HXD/EXnH motif required for Fe(II) binding, and the RXS motif for 2-oxoglutarate binding [237]. The majority of 2OGDs require Fe(II) as a cofactor and 2-oxoglutarate and molecular O_2 as cosubstrates, producing CO_2, succinate, and Fe(IV)-oxo intermediate as coproducts. The latter reacts with substrate C–H to form an alcohol concomitant with reduction of Fe(IV) to Fe(II). These enzymes have a multifunctional capacity, catalyzing several types of reactions by a single enzyme. Such reactions are hydroxylation, demethylation, sequential oxidation, ring formation, ring expansion, halogenation, epoxidation, dehydrogenation, desaturation, cyclizations, and rearrangements [235,238,239]. The reaction mechanisms are known to be like cytochrome P450s.

Previously, the capability of ring expansion by 2OGDs was reported in fungus and bacteria [231,232,240]. In general, the biological functions of 2OGDs involve either primary or secondary metabolism. Their secondary metabolic roles in plants and microbes principally involve biosynthesis or degradation/recycling of small molecules. Their primary metabolic roles include nucleic acid repair, transcriptional/protein biosynthesis regulation, lipid metabolism, hypoxia sensing, and collagen/collagen-related protein biosynthesis [241]. Several 2OG-dependent dioxygenases have previously been

functionally characterized. Such examples are *S. miltiorrhiza* 2OGD, involved in abietane type diterpenoid biosynthesis that catalyzes the hydroxylation of sugiol to produce hypargenin and crossogumerin [237].

2OGDs are divided into three classes: DOXA, DOXB, and DOXC, based on sequence similarity [73]. DOXA and DOXB are typically involved in processes like demethylation of nucleic acids and histones, proline hydroxylation during post-translation modification. Whereas the DOXC family, in which the majority of 2OGDs are included, is involved in primary and secondary metabolism. More than 100 genes for putative 2OGDs in flavonoid biosynthesis have been identified in the *A. thaliana* and *O. sativa* genomes, suggesting their structural diversity is catalyzed by various oxidative modifications [73]. Moreover, biochemical information on 2OGDs is limited because of the limited availability of their substrates, impeding our understanding of the enzymes in this superfamily.

1.8. Silvicultural effects on Thuja plicata rot resistance

Silvicultural practices have long been used to improve the quality and production of timber crops. For instance, stumping is a type of silviculture treatment in which stumps and roots are removed before plantation. The aim is to reduce root and heartwood rot disease in *T. plicata* caused by fungal pathogens i.e., *Armillaria ostoyae* and *Phellinus weirii*. Research has shown that residual tree stumps and woody debris act as a food base for the fungi to spread between stump and planted trees through rhizomorphs and eventually kill the trees [242,243]. Another silviculture strategy is species admixture i.e., plantation of disease tolerant species such as *T. plicata* with disease-susceptible species e.g., Douglas fir (*Pseudotsuga m*enziesii Mirbel Franco) [244]. Stumping has potential economic benefits for coniferous trees [245,246]. Research has found that stumping can increase forest productivity by $141/ha more than unstumped stands [245,247]. In addition, stump removal increases carbon stock in forest soils. Both stumping and tree species composition can also have an impact on the soil microbial community. Studies have shown that these treatments can promote the growth of beneficial microbes residing in soil that suppress root disease [248]. Disruption of this association between plants and microbes could have negative effects on nutrient absorption, affecting the growth and development of plants.

Silvicultural practices may affect the properties of heartwood and its secondary metabolites. However, a limited amount of research has been done to evaluate the silviculture treatment effects on *T. plicata*. Treatments such as thinning, and fertilization have no significant effect on growth and the amount of secondary metabolites in *T. plicata* [249]. In contrast to that, interesting findings were observed in a long-term research study established in a managed forest site near Salmon Arm, BC. In this study, two silviculture treatments, i.e., stumping and tree admixture, were tested. After 50 years, research has found that stump removal can reduce butt rot caused by *A. ostoyae* in *T. plicata* from 18% to 2% compared to unstumped plots. Moreover, the tree survival rate is higher in stumped plots than in unstumped plots [244]. This study also found a significant effect of these treatments on wood density; highest for monoculture *T. plicata* and lowest for *T. plicata* growing with *P. m*enziesii [7]. High wood density is a desirable attribute as it provides greater strength, stiffness as well as resistance to decay [250].

Abiotic conditions such as drought, temperature, light as well as climate change can also affect the overall outcome of these treatments [251]. For example, drought and associated higher temperatures may accelerate levels of tree mortality, either directly via decreased photosynthesis rates and cellular damage, and indirectly, via compromised ability to defend against insects and pathogens [252]. Physiological studies on several conifer species have shown that high wood density trees have a slow growth rate, narrow tracheid lumen and thus low hydraulic conductivity during drier conditions. Trees with such properties are considered more drought resistant varieties [253]. Conversely, trees with large tracheid lumen with low cell wall thickness have more hydraulic conductivity but also increase the risk for cold or drought induced embolism [254,255]. Moreover, it has been found that wood density increases with an increasing proportion of latewood and decreases with an increasing ring width [256].

1.9. Hypotheses and intended objectives

Every year, Canada generates over $1 billion from *T. plicata* trees from lumber sales [1]. Paradoxically, up to 30% of harvested volumes are lost due to extensive heartwood fungal rot, thus reducing revenues significantly [2]. Second-growth forests are nowhere near as valuable as old growth forests, producing less volume and wood that is less durable for outdoor wood products. To date, the decay in *T. plicata* is still not well understood but appear to involve an initial succession of fungal species that enter

through existing basal lesions and roots and detoxifies heartwood secondary metabolites [2,7,257,258].

Western redcedar timber has, in the past, been largely harvested from primary old-growth forests in Canada. To improve harvesting practices, there has been a momentum to adopt planting and harvesting of second-growth *T. plicata*. Second-growth *T. plicata* has been found to be susceptible to a host of fungal pathogens, such as those responsible for cedar leaf blight and heartwood rot [259]. Thus, Canada's *T. plicata* tree improvement programme is currently focused on breeding trees for improved growth, heartwood durability and resistance to disease and herbivores [260]. Previous research has shown heartwood rot resistance is largely attributed to production of antimicrobial secondary metabolites in ray cells in the sap to heartwood transition zone [3,4]. The primary heartwood secondary metabolites in this species are terpenoid tropolones and phenolic lignans [5,6]. Research has shown a weak correlation of rot resistance with the amount of tropolones in living trees. Whereas a stronger correlation has been found with lignans for wood in service [115,173]. Tropolone β-thujaplicin has free radical scavenging activity, making heartwood toxic to fungi [10]. Although its biosynthesis pathway has been defined by isotope experiment in *C. lusitanica*, the genes and enzymes involved in this pathway are unknown.

Currently, the chemical and genetic basis of heartwood rot resistance in *T. plicata* is not understood. Moreover, it takes ~20 years of growth before the main secondary metabolites, tropolones and lignans, accumulate in the heartwood [7,8]. Thus, there is a need to develop DNA markers that correlate with high levels of key secondary metabolites that can be used for early selection for heartwood rot resistance in *T. plicata*. *T. plicata* is unusual among conifers in that low to no inbreeding depression and can easily be self fertilized. Through genomic selection and early age flowering with gibberellin sprays, the breeding cycle can be reduced from 20 to 2 years [261]. Development of DNA markers would for the first time enable generational breeding for rot resistance, an issue that currently results in > $100 million losses each year.

Here we set out to identify and functionally characterize genes expressed during the heartwood formation process. While HW formation and rot resistance have been studied from a chemical perspective for the past 100 years, only a few genes with roles in these processes have been characterized in any tree species. The identification of

relevant genes and enzymatic activities proposed here may provide the target material for marker-assisted early selection breeding for rot resistance in this species. There is also the prospect of identifying the first genes and enzymes carrying out the critical expansion of the monoterpene ring into the seven-carbon ring β-thujaplicin. No such enzyme has, to date, been characterized for tropolone biosynthesis in plants. Such a discovery may also be of more general interest due to the established microbial toxicity of tropolones.

To date, the genetic basis of β-thujaplicin biosynthesis in any species is currently unknown. Based on chemical structures as well as the feeding and radiolabel tracing experiments described above, genes encoding terpinolene synthase and possibly also 3-carene and terpinyl acetate synthases are primary candidates for the first dedicated step in β-thujaplicin biosynthesis [223,229]. We generated transcriptomes from different wood fractions of *T. plicata* trees and identified two contiguous sequences encoding putative monoterpene synthases that are highly upregulated in the sapwood to heartwood transition zone in which secondary heartwood metabolites are produced. We described three cDNAs encoding three putative diterpene synthases that are highly upregulated in the SW (Sapwood) to HW (Heartwood) transition zone. Here we are using gene expression during heartwood formation as a proxy of heartwood resistance.

We tested these hypotheses by comparing the gene expression levels between stumped and unstumped and different tree mixtures at two different time points, June, and August.

This research is based on the following hypotheses:

I. Chapter 2: One or several genes expressed where heartwood is formed and that encode putative monoterpene synthases produce the first dedicated precursor to β-thujaplicin, terpinolene.

II. Chapter 3: Three Diterpene synthases expressed in the heartwood to sapwood transition zone are monofunctional enzymes and synthesize hitherto unknown diterpenes in the *T. plicata* heartwood.

III. Chapter 4: One or more identified dioxygenases catalyze the ring expansion resulting in β-thujaplicin.

IV. Chapter 5: A 50-year-old silvicultural treatment results in gene expression differences to this day, and RNA-seq can be used to shed light on potential explanations for that.

The research demonstrates the following significance:

I. Chapter 2: First genetic and biochemical evidence for terpinolene synthase and 3-carene/terpinolene synthase activities in *T. plicata* heartwood formation, providing candidates for first dedicated step in β-thujaplicin biosynthesis, the most abundant tropolone in *T. plicata* heartwood.

II. Chapter 3: To date, no gene has been identified in *T. plicata* with a potential role in the biosynthesis of diterpenes during heartwood formation. We also characterized their enzymatic activities, revealing hitherto unknown functions in this species. The identity and functions of diterpenes in *T. plicata* heartwood are unknown, as are the genes behind their biosynthesis. These genes and their enzyme products add to a picture of heartwood rot resistance hitherto focused entirely on tropolones and lignans in *T. plicata*.

III. Chapter 4: The genetic and molecular basis of tropolone biosynthesis is unknown in any plant species. The discovery of a gene and its enzyme carrying out the ring expansion characteristic of a plant tropolone would be a novelty, provide an avenue to similar functions in other plant species and provide an obvious target(s) for selection for rot resistance in *T. plicata*.

IV. Chapter 5: First gene expression study on mature *T. plicata* growing in natural forest. The results shed light on how a 50-year-old treatment can affect wood quality to this day and may influence the selection of sites at which similar treatments should be done to improve yield and wood quality.

Chapter 2. Identification of two monoterpene synthases and one sesquiterpene synthase expressed during heartwood formation in Western redcedar (*Thuja plicata*) trees

2.1. Materials and Methods

2.1.1. Plant Material and RNA isolation

Stem sections of approximately 60-year-old *T. plicata* trees in the forest of Simon Fraser University campus were cut with a chain saw, frozen in liquid nitrogen and stored at −80°C. Samples were cut with chisel and hammer into following fractions: bark to cambium including phloem and excluding most dead outer bark, ~2 cm outer sapwood, ~2 cm inner sapwood, and transition zone consisting of two annual rings of sapwood and two rings of heartwood. Sticks of samples were freeze-dried for 3 days, milled to ≤0.5 mm particle size. To recover low concentration of RNA, approximately 1.5 g powder was rehydrated in ice-cold extraction buffer (2% CTAB; 200 mM Tris, pH 8.0; 50 mM EDTA, pH 8.0; 1 M NaCl; 0.5% activated charcoal; 1.5% polyvinylpolypyrrolidone; and 1.5% β-mercaptoethanol), ground in liquid nitrogen and transferred to Oakridge centrifuge tubes. After evaporation of nitrogen, 10 mL pre-warmed (65°C) extraction buffer was added, mixed, and tubes were incubated in a water bath (65°C) for 10 min. The homogenate was extracted twice with an equal volume of chloroform (13,000 Relative Centrifugal Force (RCF) at the bottom). The aqueous phase was mixed with an equal volume of 100% isopropanol and kept at −20°C for 2–3 h to precipitate nucleic acids. After centrifugation at 17,000 RCF at the bottom of the tubes, the pellet was resolubilized in 0.5 mL SSTE (1 M NaCl; 0.5% SDS; 10 mM Tris-HCl, pH 8.0; and 1 mM EDTA, pH 8.0) buffer followed by chloroform extraction. The RNA was precipitated overnight on ice with ¼ volume of 10 M LiCl at 4°C in 1.5 mL tubes. After 30 min centrifugation at 14,000 RCF at 4°C, the pellet was washed in 75% ice-cold ethanol, dried, and resuspended in diethylpyrocarbonate (DEPC)-treated water. RNA integrity was analyzed by bleach gel electrophoresis [262]. Residual DNA was removed, and first strand cDNA synthesized (AccuRT genomic DNA removal, EasyScriptTM cDNA synthesis kit, abmgood.com).

2.1.2. Selection of candidate genes

An RNA-seq analysis identified an estimated 50,000 different transcripts that are expressed in *T. plicata* wood. Among them, approximately 800 are upregulated in the sapwood to heartwood transition zone [263]. From these 800 transcripts, 20 different putative TPS-encoding genes were identified, based on sequence similarity to known conifer terpene ORFs. Among them, three transcripts encoding full-length putative *T. plicata* monoterpene synthases (TpTS1 and TpCS1) and one sesquiterpene synthase (TpES1) were selected. The de-novo assembled contiguous sequences were confirmed by polymerase chain reaction (PCR) amplification from cDNA, cloning, and Sanger sequencing. The accession numbers were for TpTS1 and TpCS1, are OM324392 and OM324393, respectively. BLASTX and BLASTP were used to assess sequence similarity to other known TPS sequences. Multiple alignment of protein sequences was conducted by UniProt Align tool [264] and the conserved motifs were visualized by Jalview software [265]. The plastid transit peptide in the ORFs of the selected TPS genes was identified with iPSORT and TargetP1.1 [266,267] and deleted during plasmid cloning.

2.1.3. Molecular cloning of mono- and sesqui TPS genes

To isolate the open reading frame (ORF) of mono- and sesqui TPS cDNAs two sets of primers were designed for PCR amplification (Table B.2). The online Expasy translate tool was used to double-check the continuity of the ORF-His tag fusion protein. A nested polymerase chain reaction (PCR) was performed to amplify three predicted full length TPS cDNAs with the first set of primers according to CloneAmp™ HiFi PCR premix (Takara Bio, Kusatsu, Japan) protocol with a 59°C annealing temperature. PCR products were diluted 100-fold in water and used for a second round of PCR (60°C annealing temperature) using a second set of primers designed to delete the transit peptide and generate an N-terminal fusion to a His tag after cloning. Resulting single band products were cloned into *Ssp*I (New England Biolabs, Ipswich, USA) digested pETHis6GST expression vector (Addgene, Watertown, MA, USA) using the In-Fusion HD Cloning kit (Takara Bio Inc., San Jose, CA, USA). To verify the full-length sequence via sequencing, construct was first transformed into Stellar chemically competent cells (Takara Bio, Inc.). Colonies harbouring plasmid with insert were identified by colony

PCR. Plasmids were purified and inserts were Sanger-sequenced from both directions. Sequences contained an ORF with N-terminal His tag followed by in-frame fusion to the predicted original mono- and sesquiterpene synthase peptide.

2.1.4. Prokaryotic expression and purification of recombinant TPS enzymes

Recombinant protein expression and purification were performed as described [146]. For the overexpression of putative terpene synthase genes, the ORF-containing vector was transformed into BL21 StarTM (DE3, ThermoFisher, Waltham, MA, USA) chemically competent *E. coli* cells. To functionally express TPS genes, four individual colonies of transformants were inoculated in 5 mL Luria-Bertani (LB) medium supplemented with 35 µg/mL Kanamycin and cultured overnight at 37°C, at 225 rpm. The following day, bacterial cell suspension was diluted 1:50 with Terrific Broth medium with the same concentration of antibiotic. Incubation was continued at 37°C until an OD_{600nm} reached 0.9. Cultures were then cooled on ice for 20 min and protein expression induced by the addition of IPTG to a final concentration of 0.2 mM. The incubating temperature was lowered to 16°C and grown for another 19 h at 220 rpm. Cells were harvested by centrifugation at 4°C, 5000 RCF for 30 min and pellets were used immediately or stored at - 80°C. Cell pellets were resuspended and disrupted by sonication in 3 mL fresh lysis buffer (20 mM NaPO4, 500 mM NaCl, 30 mM imidazole, 1 mg/mL DNase, 1 mg/mL RNase, 100 mM $MgCl_2$, 200 mM PMSF, and 25 mg lysozyme, pH 7.4) for 3x15 s, 30 s off between, 30% strength followed by another 3x15 s sonication. The cleared lysates were collected by centrifugation (30 min, 12,000 RCF, 4°C). The supernatant was removed and loaded onto His Spin Trap Ni-affinity columns (GE Healthcare, Chicago, IL, USA) to purify His-tagged recombinant protein following the manufacturer's protocol. Purified enzymes were desalted into buffer (25 mM HEPES pH 7.2, 100 mM KCl, and 10% glycerol) at 4°C, using PD Mini-Trap G-25 desalting columns (GE Healthcare). Protein concentrations were determined on a nanodrop instrument (ThermoFisher, Waltham, MA, USA) and assayed immediately for enzyme activity. Purified protein samples were also examined on a pre-cast 10% SDS-polyacrylamide gel (data not shown, Bio-Rad Laboratories Inc., Montreal, QC, Canada)

2.1.5. In Vitro Enzyme assays

Standard enzyme assays were carried out in a total volume of 250 µl with approximately 50 µg of purified protein, buffer (25 mM HEPES, pH 7.2, 100 mM KCl, 10 mM $MnCl_2$, 5 mM fresh DTT, 10% glycerol) and 8 µg GDP for monoterpene synthase assay; (25 mM HEPES, pH 7.2, 10 mM $MgCl_2$, 5 mM DTT, 10% glycerol, and 8 µg FDP ((E, E)-farnesyl diphosphate) for sesquiterpene synthase assay; (50 mM HEPES pH 7.2, 100 mM KCl, 7.5 mM $MgCl_2$, 20 µM $MnCl_2$, 5 mM fresh DTT, 5% glycerol, and 8 µg GGDP ((E,E,E)- geranylgeranyl diphosphate) for diterpene synthase assay. All substrates were purchased from Sigma-Aldrich (Oakville, ON, Canada). The mixture was overlaid with 250 µl of pentane, with addition of 10 ppm IBB (Isobutyl benzene) as internal standard, to trap terpene products during the assays, and incubated at 30°C for 1 h. The reaction was stopped by vortexing for 30 s to extract terpene products into pentane layer. To separate the phases, samples were frozen at -80°C, thawed and centrifuged for 30 min at 1000 RCF at 4°C. Controls were performed under the same condition using the purified product from empty expression vector without any TPS cDNA and tested for their product formation.

2.1.6. Product identification by GC-MS analysis

GC-MS analysis was performed on an Agilent 6890 gas chromatograph coupled to an 7683B series autosampler and 5975 Inert XL Mass Selective Detector (Agilent Technologies, Santa Clara, CA, USA). Compounds were separated on an Agilent DB1MS column (30 m x 0.25 mm ID, 0.25 µm film) at 35 cm/s with a He flow of 0.9 mL/min. The Inlet was operated splitless at a temperature of 250°C, pulse pressure 40 psi. The following temperature program was used: initial temperature of 40°C, 3 min hold, and 15°C /min ramp to 300°C, 5 min hold. Compounds were identified by comparing retention times and mass spectra with those of authentic standards or Wiley125 MS-library (Wiley Publishing, Hoboken, NJ) matches. The retention indices of the identified compounds are provided in Table B.3. The authentic standards were purchased from Sigma-Aldrich. The identity of elemol which was confirmed by running elemi essential oil that produces elemol as one of the constituents of the oil. The elemi essential oil was purchased from Aromatics Canada Inc.

2.2. Results

2.2.1. Identification of Thuja plicata mono- and sesquiterpene synthase genes

From a *T. plicata* stem transcriptome, three contiguous sequences encoding putative mono- and sesquiterpene were retrieved by using other conifer terpene synthase cDNAs as queries in BLASTX searches (Figure A.1). Based on results presented below, the sequences were named *Thuja plicata* Terpinolene synthase 1 (TpTS1), *Thuja plicata* Carene Synthase1 (TpCS1) and *Thuja plicata* Elemol synthase 1 (TpES1). Deduced amino acid sequences of the open reading frames of TpTS1, TpCS1 and TpES1, resembled most closely known conifer mono- and sesquiTPSs. The contiguous cDNA sequences of TpTS1, TpCS1 and TpES1 were 2108bp, 2057bp and 1993bp in length, along with predicted open reading frames encoding 614, 615 and 569 amino acids residue proteins, respectively. The molecular masses are predicted to be 70.83, 70.47 and 65.46 kDa, respectively. The full-length transcript sequences were confirmed by polymerase chain reaction (PCR) amplification from cDNA, cloning, and Sanger sequencing. The sequences of TpTS1 and TpCS1 were deposited in GenBank and given the accession numbers OM324392 (TpTS1) and OM324393 (TpCS1). A search of NCBI GenBank using the predicted TpTS1 and TpCS1 protein sequences as queries identified the highest sequence similarity hits with monoterpene synthases from several conifer species with percentage identities ranging from 47 to 71% (Table 2.1).

Table 2.1. Sequence similarity between TpTS1 and TpCS1 and top (5) hits in NCBI Genbank.

Species	Protein activity	Accession #	Similarity to TpTS1 (%)	Similarity to TpCS1 (%)
Chamaecyparis formosensis	α-pinene synthase	EU099434.1	52.36%	71.90%
Abies grandis	Myrcene synthase	O24474.1	50.71%	53.66%
Thuja plicata	Sabinene synthase	KC767281.1	49.84%	54.78%
Picea abies	(+)-3-carene synthase	Q84SM8	49.74%	52.15%
Pseudotsuga menziesii	terpinolene synthase	AAX07264.1	47.89%	50.87%

TpES1 showed the highest sequence similarity with sesquiterpene synthases, with 34.92% similarity to *Chamaecyparis formosensis* β-cadinene synthase and 41.69% similarity to *Abies grandis* γ-humulene synthase. A phylogenetic analysis, including conifer mono-, sesqui- and diterpene synthases showed that TpTS1 and TpCS1 clustered with monoterpene synthases (Figure 2.1). Among conifer monoterpene synthases, TpTS1 and TpCS1 separated from the Pinaceae family of conifers to form a clade with a sabinene synthase, TpSS1, identified from the same species [182] and Cf-Pin, from *Chamaecyparis formosensis,* which is also a Cuppressaceae species [268]. TpES1 formed a separate clade with sesquiterpene synthases.

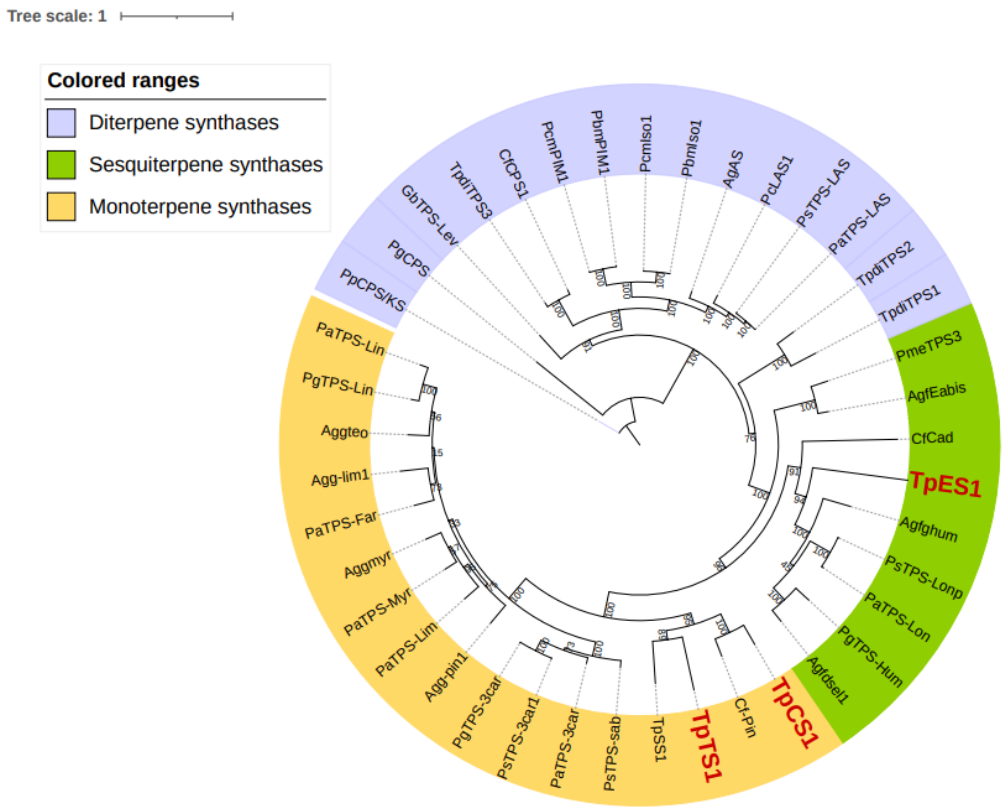


Figure 2.1. Phylogenetic tree of conifer mono-, sesqui- and diterpene synthases.

Phylogenetic analysis was performed with 39 full-length mono, sesqui- and diterpene synthases. Proteins identified in this study are highlighted in red. TPS (Terpene synthases) of the Cupressaceae cluster apart from TPS of the Pinaceae, indicative of the independent diversification and evolution of specific TPS functions in the Cupressaceae and the Pinaceae. The tree is rooted with bifunctional diterpene synthase *ent*-CPS/*ent*-kaurene from the moss species *Physcomitrella patens* (PpCPS/KS). Numbers indicate bootstrap support of 100

repetitions. The full abbreviation of the protein sequences and corresponding accession numbers are listed in Table B.1.

2.2.2. Predicted proteins of TpTS1, TpCS1 and TpES1 harbour the functional motifs typical of terpene synthases

When the predicted open reading frames (ORFs) of the TpTS1, TpCS1 proteins were aligned with related conifer monoTPS ORF, the predicted conserved N-terminal RRX8W motif and the C-terminal aspartate-rich DDXXD motif aligned well, showing that they are present and intact in the TpTS1 and TpCS1 ORFs (Figure 2.2). In addition, analysis using prediction software (see materials and methods) indicated the presence of putative N-terminal chloroplast transit peptides in the TpTS1 and TpCS1 ORFs. Taken together, the predicted TpTS1 and TpCS1 proteins contain all three of the functional motifs expected of functional monoTPSs. The alignment of the predicted TpES1 protein with conifer mono- and sesquiTPSs revealed the presence of C-terminal DDXXD motif but a defective RRX_8W motif. TpES1 also lack a transit peptide, indicating that it may localize to the cytosol, consistent with sesquiTPSs from various species.

Protein	Transit peptide	R	R				X8				W	D	D	X	X	D
TpTS1	MALSSAFSYNFCLKSQPSQSQSQSQTQSHT	R	R	I	G	N	H	H	P	N	L	W	D	D	T	Y	D
TpCS1	MAVSCITPLASTMVGQKLPLGRVPSIQRKP	R	R	I	G	N	H	H	P	N	L	W	D	D	I	Y	D
TpSS1	MALFSASTSVLSSCLKSPPNHHVKLFNKNS	R	R	T	G	N	H	H	P	N	L	W	D	D	I	Y	D
Cf-Pin	MSLGCITPLASAMVGPKLVRPLIHHNPLFH	R	R	I	G	N	H	H	P	N	L	W	D	D	M	Y	D
PaJF67	MSVISILPLASKSCLYKSLMSSTHELKALC	R	R	I	G	D	H	H	S	N	L	W	D	D	M	Y	D
PmeTPS2	MSLISMAPLAPKSCLHKPFIGSTHEPKVFC	R	R	I	G	D	Y	H	S	N	I	W	D	D	M	Y	D
TpES1	.....	M	E	M	P	E	K	S	L	N	L	W	D	D	L	F	D
Cf-Cad	.....	R	R	V	A	E	F	H	P	N	V	W	D	D	I	F	D
Agfghum	.....	S	I	T	S	N	R	H	G	N	M	W	D	D	L	Y	D

Figure 2.2. Putative catalytic motifs in TpTS1, TpCS1 and TpES1.

Amino acid alignments of *T. plicata* TpTS1, TpCS1 and TpES1 with *Pseudotsuga menziesii* terpinolene synthase (PmeTPS2, AAX07264.1), *Picea abies* (+)-3-carene synthase (PaJF67, Q84SM8.1), *T. plicata* sabinene synthase (TpSS1, AGO02736.1), *Chamaecyparis formosensis* α-pinene synthase (Cf-Pin, C3RSF5.1), *Chamaecyparis formosensis* β-cadinene synthase (Cf-Cad, I1ZHA5.1), *Abies grandis* γ-humulene synthase (Agfghum, O64405.1). Full-length alignments are shown in Figure A.2.

2.2.3. TpTS1, TpCS1 and TpES1 are expressed at elevated levels in the sap to heartwood transition zone

Differential expression analyses of RNA-seq results showed that transcript levels of both *TpTS1* and *TpCS1* and *TpES1* were elevated in the sap to heartwood transition zone (TZ), relative to bark (B), outer sapwood (OS) and inner sapwood (IS) (Figure 2.3). The transcript levels of *TpTS1*, *TpCS1* and *TpES1* were approximately 16, 166 and 191 times higher in the TZ relative to the other fractions of the stem. The statistical analyses showed a significant difference in expression for *TpCS1* (p value <0.0002) and *TpES1* (p value <0.035). Although, the expression of *TpTS1* was higher in transition zone, the expression difference relative to other fractions was not significant. Taken together, these results demonstrate these putative terpene synthases are upregulated in the transition zone, and that their expression is higher than the average for genes expressed in the transition zone.

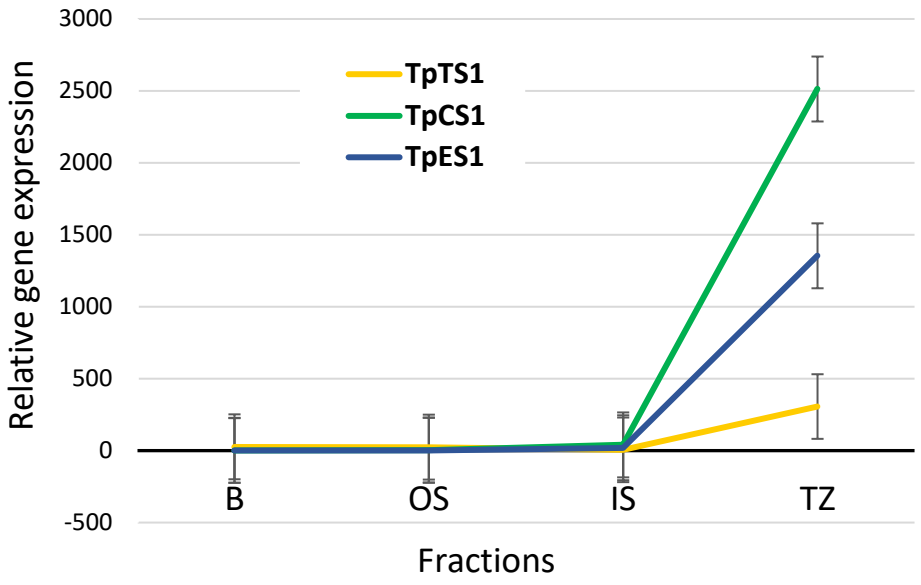


Figure 2.3. TpTS1, TpCS1 and TpES1 are expressed at elevated levels in the sap to heartwood transition zone.

Expression of *TpTS1*, *TpCS1* and *TpES1* in RNA-seq libraries derived from four different fractions of *T. plicata* wood. B-bark, OS-outer sapwood, IS-inner sapwood, TZ-sapwood to heartwood transition zone. Standard error bars were calculated based on two biological

replicates. TMM (trimmed mean of M-values) normalized TPM (transcripts per million) values were used to calculate relative expression levels. Statistical differences in different fractions were determined in JMP 15.0.0 by two-way ANOVA and Tukey HSD tests.

2.2.4. TpTS1 produces terpinolene and TpCS1 produces 3-carene and terpinolene from geranyl diphosphate

To determine the function of TpTS1 and TpCS1, cDNAs were cloned into an expression plasmid, protein produced in *E. coli* cells, and Histidine-tagged recombinant proteins were purified. Enzyme assays using the TpTS1 and TpCS1 proteins and the substrates for sesqui- and diterpene synthases, i.e., farnesyl diphosphate (FDP) and geranyl geranyl diphosphate (GGDP), yielded no product (data not shown). In contrast, incubation of TpTS1 protein with the monoterpene substrate geranyl diphosphate (GDP) resulted in a single major product with a gas chromatography-mass spectrometer (GC-MS) retention time and mass spectrum identical to a terpinolene standard (Figure 2.4). This gene and protein were therefore named *T. plicata* terpinolene synthase 1 (TpTS1). Assays with the TpCS1 protein and GDP resulted in GC peaks and mass spectra identical to 3-carene (80% of total peaks) along with terpinolene (14% of total peaks), α-pinene (3% of total peaks), α-fenchene (0.6% of total peaks), myrcene (0.2% of total peaks) (Figure 2.5). For TpCS1, the determination of above percentage is relative assuming the response factor [269] for each peak is same. Terpene identification was further validated by comparing retention times and mass spectra with a library. Extracts prepared from the same *E. coli* strain transformed with a plasmid lacking cDNA insert did not yield any detectable monoterpene (data not shown).

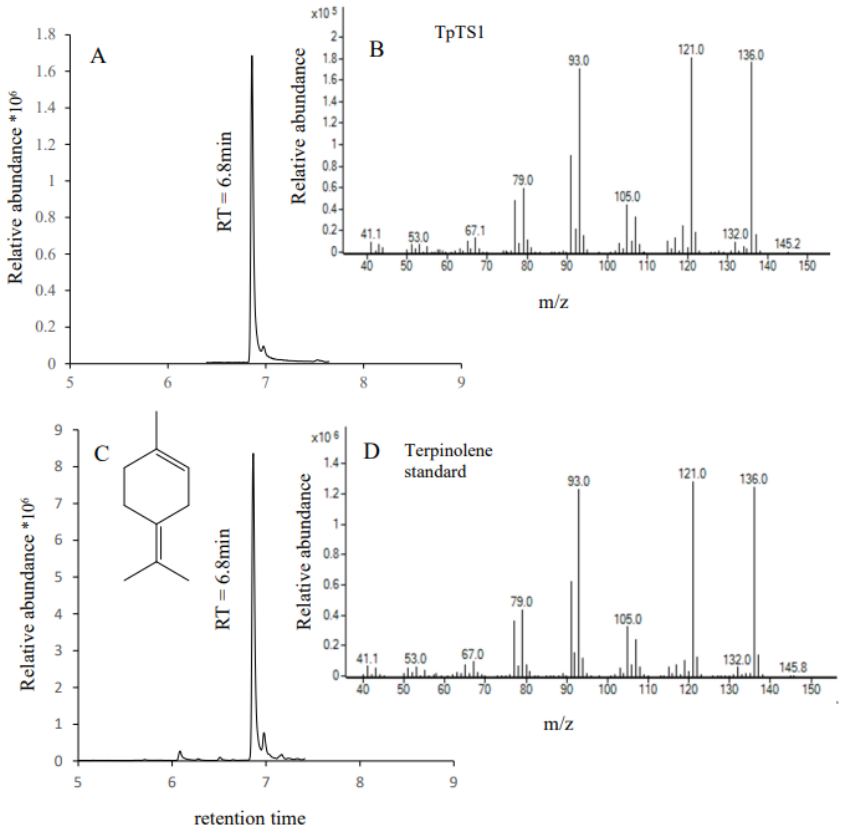


Figure 2.4. TpTS1 incubation with geranyl diphosphate (GDP) results in the production of terpinolene.

GC chromatograms of product from assay with TpTS1 (A) and authentic terpinolene standard (C) showing identical retention times of their peaks. Total ion mass chromatogram of TpTS1 product (B) and an authentic standard of terpinolene (D) showing nearly identical mass spectra.

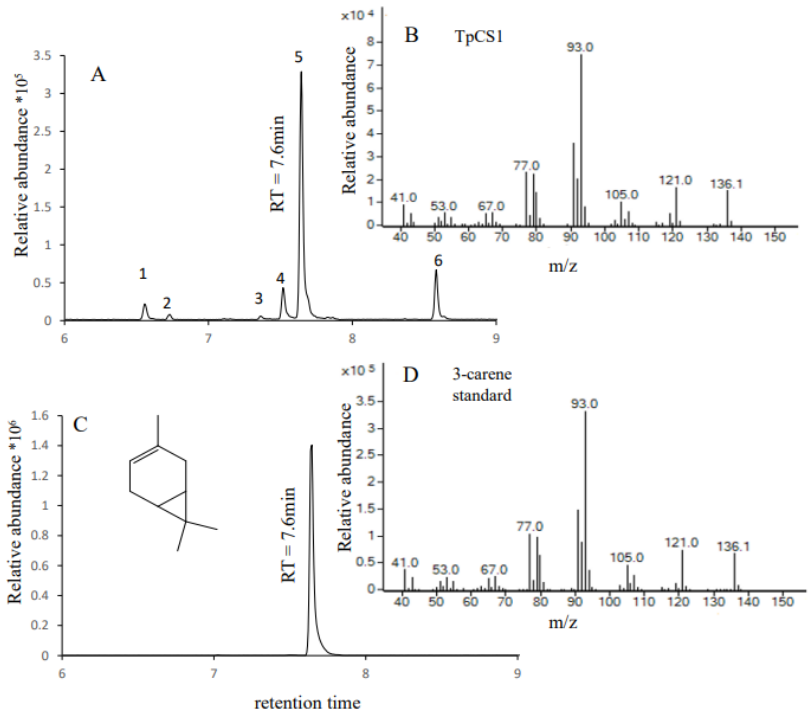


Figure 2.5. GC-MS analysis identifies 3-carene and terpinolene as the major products from incubation of TpCS1 with geranyl diphosphate (GDP).

GC chromatograms of product from assay with TpCS1 (A) and authentic 3-carene standard (C) showing identical retention times of the most abundant TpCS1 product and 3-carene (peak 5). Total ion mass chromatogram of TpTS1 product 5 (B) and an authentic 3-carene standard (D) showing nearly identical mass spectra. Additional peaks in (A) correspond to α-pinene (1), α-fenchene (2), myrcene (3), isobutyl benzene internal standard (4), terpinolene (6) based on comparison of mass spectra with authentic standards (1, 3, 6) and library spectrum (2) (Figure A.3).

2.2.5. TpES1 produces the sesquiterpene elemol from farnesyl diphosphate

Enzyme assays using the TpES1 protein and the substrates for mono- and diterpene synthases, i.e., geranyl diphosphate (GDP) and geranyl geranyl diphosphate (GGDP), yielded no product (data not shown). In contrast, incubation of TpES1 protein with the sesquiterpene substrate FDP resulted in a single major product with a gas chromatography (GC) retention time and mass spectrum identical to an elemol standard (Figure 2.6). This gene and protein were therefore named *T. plicata* Elemol Synthase 1 (TpES1).

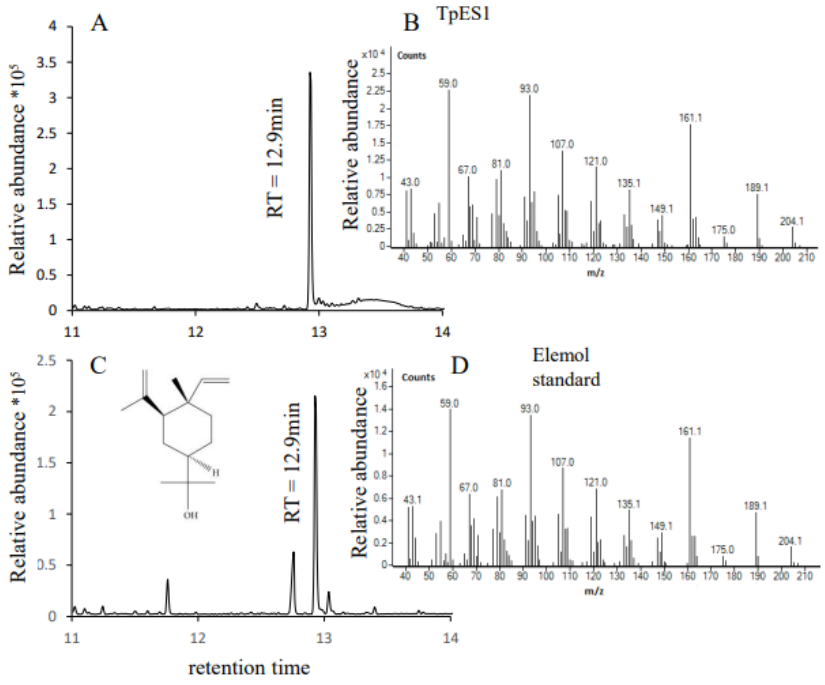


Figure 2.6. GC-MS analysis identifies elemol as the single product from incubation of TpES1 with farnesyl diphosphate (FDP).

GC retention times of elemol produced by TpTS1 (A) = total ion chromatogram (TIC), (B) = mass spectrum of elemol, (C, D) = TIC and mass spectrum of an authentic standard of elemol. The authentic standard of elemol was obtained by running elemi essential oil that produces elemol as one of the constituents of the oil.

.

2.3. Discussion

β-thujaplicin is an abundant wood tropolone in many Cupressaceae species. [61,224,229,270,271]. Given their common ancestry, these species are likely to use similar pathways for β-thujaplicin biosynthesis. We and others have an interest in understanding the biochemical and genetic basis of β-thujaplicin biosynthesis, for the development of marker-assisted selection to improve heartwood rot resistance, and for the potential production of this powerful antimicrobial compound in alternative hosts [260,261,272].

The genetic basis of β-thujaplicin biosynthesis is to date completely unknown in any plant species. In *C. lusitanica* cell cultures, however, the biosynthetic pathway was partially resolved by feeding experiments and microsomal assays [195,223]. Based on isotope tracer experiment in *C. lusitanica*, terpinolene was part of the predicted first biosynthesis step dedicated to the production of the heartwood tropolone β-thujaplicin [223]. Prior to that, another monoterpene 3-carene was theoretically proposed as an intermediate of β-thujaplicin biosynthesis by the idea of organic chemistry. Feedings of *C. lusitanica* suspension cultures with 3-carene significantly improved the β-thujaplicin production of *C. lusitanica* suspension cultures, indicating its possible involvement in β-thujaplicin biosynthesis [229].

In this work we identified two genes that are upregulated during heartwood formation. The encoded proteins of one gene produced terpinolene as a single major product whereas the other produced both 3-carene and terpinolene in *in vitro* enzyme assays. Based on the above chemical evidence from *C. lusitanica* cell cultures, these genes are therefore likely to play a role in β-thujaplicin tropolone production in *T. plicata*. Terpinolene and 3-carene synthase-encoding genes have been identified in other plant species [273], active in sapling stems [153], shoot and bark [274] and seedlings [150]. However, no heartwood specific terpinolene and 3-carene synthase have been identified.

Here, we also identified a full-length cDNA sequence which was upregulated during heartwood formation and has sequence similarity to those of sesquiterpene synthase. Enzyme assay of purified recombinant protein produced sesquiterpene alcohol elemol, confirming its identity as elemol synthase (TpES1). When comparing the conserved motifs required for catalytic functions, TpTS1 and TpCS1 retained the N-terminal RRX_8W motif required for the ionization and isomerization of the substrate [159]. While TpES1 had a substitution in the two conserved arginine (R) nucleotides and was found as MEX_8W [159] similar to sesquiterpene synthase (i.e., γ-humulene synthase) in *Abies grandis* [179]. All three proteins retained the other C-terminal DDXXD motif which is required for the binding of divalent metal ion cofactors such as Mg^+ or Mn^+ to initiate binding and activation of the diphosphate moiety of the substrate [144,167]. TpTS1 and TPCS1 contained an N-terminal transit peptide required for plastid import, consistent with other monoterpene synthases [144,167]. While the absence of transit

peptide in TpES1 indicates its localization in cytosol where sesquiterpene biosynthesis takes place.

Research has shown that terpenoid based defense response can be induced by phytohormone treatment [275]. Such as, the enzyme activity levels of the (+)-3-carene synthase and the corresponding accumulation of (+)-3-carene can be induced to a higher fold change than any other TPS or metabolite measured [276]. Expression of sesquiterpene gene elemol synthase in *S. tuberosum* leaves can be induced in response to bacterial and fungal infection as well as methyl jasmonate treatment [277]. In our study, we also attempted to inject methyl jamonate and ethephon (precursor of ethylene) in mature *T. plicata* trees but didn't see any significant upregulation of terpene synthase genes (Data not shown).

The toxicity of the identified monoterpenes and sesquiterpene is known in several species. Such as, in *P. sitchensis*, 3-carene is associated with resistance to pine weevil (*Pissodes strobi*), where resistant trees produced significantly higher levels of (+)-3-carene whereas susceptible trees produced trace amounts of (+)-3-carene [278,279]. Two monoterpenes, terpinolene and (+)-3-carene, were associated with resistance in genotypes originating from a high weevil hazard area [280]. Terpinolene has shown fumigant toxicity against *L. Bostrychophila* [281]. Sesquiterpene β-elemol is one of the major constituents of Japanese cedar (*Cryptomeria japonica* D. Don) leaf essential oil. Studies on this species had shown that β-elemol has antitermite activity against the subterranean termite *Coptotermes formosanus* Shiraki, responsible for most wood destruction in countries such as Taiwan, Japan, and parts of the United States [282]. Elemol and eudesmol extracted from *C. japonica* leaf essential oil also exhibited antifungal activity against dermatophyte *Trichophyton rubrum* [283].

Based on the above mentioned strong antifungal properties as well as the high levels of these compounds in heartwood, these genes are likely to play a major role in rot resistance of growing *T. plicata* trees. To our knowledge, these are the first identified terpinolene, 3-carene and elemol synthase-encoding genes, active during heartwood formation. These genes are also not identified in any Cupressaceae species. The identified genes provide prime candidates for markers-assisted selection and study of genetic variation that affects protein function and expression in *T. plicata* tress.

Chapter 3. Identification of Three Monofunctional Diterpene Synthases with Specific Enzyme Activities Expressed during Heartwood Formation in Western Redcedar (*Thuja plicata*) Trees

3.1. Materials and Methods

3.1.1. Plant materials

RNA was extracted from approximately 60-year-old trees taken down at the Simon Fraser University campus. Cut sections were frozen in liquid nitrogen and radial fractions were taken using chisel and hammer. The following fractions were taken; bark to cambium excluding most dead outer bark, ~2 cm sapwood, ~2 cm older sapwood, and transition zone consisting of two annual rings of sapwood and two rings of heartwood. Sticks of samples were freeze-dried for 3 days, milled to ≤0.5 mm particle size. RNA was prepared by following the method as described in Chapter 1. RNA integrity was analyzed by bleach gel electrophoresis [262]. Residual DNA was removed, and 1st strand cDNA was synthesized (AccuRT genomic DNA removal, EasyScriptTM cDNA synthesis kit, abmgood.com).

3.1.2. Selection of Candidate Genes

Three contiguous cDNA sequences encoding putative *Thuja plicata* diterpene synthases (TpdiTPS1-3) were selected based on expression in an RNA-seq library made from the transition zone where heartwood was formed [263]. The transcript sequences were confirmed by polymerase chain reaction (PCR) amplification from cDNA, cloning, and Sanger sequencing. The accession numbers were for TpdiTPS1-3, MT468207, MT468208, and MT468209, respectively. BLASTX and BLASTP (NCBI, National Center for Biotechnology Information) were used to assess sequence similarity to other known diTPS sequences and the presence of conserved RRX_8W, DXDD, and DDXXD motifs. The protein sequences were aligned by the Clustal Omega program and formatted in Gene Doc. Chloroplast transit peptide in the ORFs (Open Reading Frames)

of the selected TPS genes was identified with iPSORT and TargetP1.1. and deleted during plasmid cloning.

3.1.3. Cloning of diTPS cDNAs

Two sets of primers were designed to amplify the full-length cDNA (Table B.4), keeping in mind that the vector would be linearized with *SspI* restriction enzyme (New England Biolabs, MA, USA). The online Expasy translate tool was used to double-check the continuity of the ORF-His tag fusion protein. A nested polymerase chain reaction (PCR) was performed to amplify three predicted full-length TPS cDNAs with the first set of primers according to CloneAmp™ HiFi PCR premix (Takara Bio, CA, USA) protocol with a 55°C annealing temperature. PCR products were diluted 10-fold in water and used for PCR using a second set of primers and 60°C annealing temperature. Resulting single-band products were cloned into *SspI*-digested pETHis6GST expression vector (Addgene) by the Sequence and Ligation Independent Cloning method (SLIC; [284]). The mix was transformed into Stellar chemically competent cells (Takara Bio, CA, USA). Colonies with recombined plasmid were identified by colony PCR, used for plasmid purification. Inserted fragments were fully sequenced by Sanger, confirming expected nucleotide sequences. All resulting recombinant proteins contained the open reading frame with an N-terminal His tag.

3.1.4. Transformation and In Vivo Coexpression

Recombinant protein expression, purification, and assays were performed as described [285]. For the expression of putative terpene synthase genes, plasmid carrying the gene of interest together with individual pGG, pGGeC, pGGsC, pGGnC plasmids was sequentially transformed into BL21 Star™ (DE3, ThermoFisher, CA, USA) chemically competent *E. coli* cells. To confirm the stereochemistry of normal-CPS activity, TpdiTPS3 was co-expressed with pGG and a class I labdane related diterpene synthase, i.e., *Arabidopsis thaliana* kaurene synthase (AtKS) that only reacts with *ent*-CDP. AtKS, cloned in pDEST15/rAtKS [285], was also co-expressed individually with pGGec and pGGnc to confirm the presence and absence of *ent*-kaur-16-ene. To express TPS proteins, four independent colonies of transformants were grown in 5 mL Luria-Bertani medium supplemented with 35 µg/mL kanamycin and 25 µg/mL chloramphenicol and cultured overnight at 37°C, at 225 rpm. The following day, the

bacterial cell suspensions were diluted 1:50 with Terrific Broth medium containing the same antibiotics and grown as above until an OD_{600} nm of 0.6. Cultures were then shifted to 16°C for 1 h before induction with 0.5 mM IPTG (isopropyl-1-thio-beta-D-galactopyranoside) and 5 g/L glycerol and grown for ~72 h at 16°C, at 220 rpm.

3.1.5. Isolation of Recombinant Protein Products

The bacterial cells were collected by centrifugation at 1400 g for 30 min. The hydrocarbon product was extracted from the media with an equal volume of hexane with 2% (v/v) ethanol [286,287].

3.1.6. Silica and Alumina Chromatography

A Pasteur pipette was loaded with glass wool plug, ~0.6 g silica gel, ~0.1 g $MgSO_4$, and washed five times with hexane. The organic phase (8 mL) was passed over the silica gel column. In the case of oxygen-containing products, the media was extracted with an equal volume of hexane: ethyl acetate (6:1), and the organic phase was passed through a Pasteur pipette with glass wool plug and ~0.6 g basic alumina column overlaid with ~0.1 g $(NH_4)_2SO_4$. Approximately 8 mL of the eluate was evaporated under an N_2 stream. The sample was resuspended in 500 µL of hexane and derivatized overnight at room temperature with N,O-bis(trimethylsilyl)trifluoroacetamide (MilliporeSigma) before GC-MS analyses.

3.1.7. GC-MS Analysis

GC-MS analysis was performed on an Agilent 7890B gas chromatograph and a 5977A MS Detector at 70 eV (Agilent Technologies Inc., CA, USA). One microliter of the extract was injected, and compounds were separated on a DB5 column (30 m × 0.25 mm ID, 0.25 µm film) with an He flow of 1 mL/min. The inlet was operated at 250°C, splitless mode. The GC temperature program was as follows: 50°C, hold 5 min, and 10 °C/min ramp to 280°C, 10 min hold. Diterpenes were identified by comparing retention time and mass spectra with standards kindly provided by Drs. Reuben J. Peters and Meimei Xu, Iowa State University, USA. The retention indices of the identified compounds are provided in Table B.5.

3.2. Results

3.2.1. Identification of Three Putative Diterpene Synthases Expressed in Thuja plicata Sap to Heartwood Transition Zone

To identify putative diterpene synthases, we searched a transcriptome derived from *T. plicata* tree stem tissues with a putative diterpene synthase annotation. We identified three contiguous cDNA (complementary DNA) sequences (Figure A.4) and named them temporarily TpdiTPS1–3. The contiguous cDNA sequences of TpdiTPS1, 2, and 3 were 2731 bp, 2898 bp, 3080 bp, along with predicted open reading frames encoding 846, 879, and 865 amino acid residue proteins, respectively. The contigs had extensive read coverage at 5′ and 3′ ends (data not shown). The transcript sequences were confirmed by polymerase chain reaction (PCR) amplification from cDNA, cloning, and Sanger sequencing. A comparison of predicted TpdiTPS1, 2, and 3 protein sequences with that of known protein sequences showed a maximum of 74%, 80%, and 87% identity, respectively, across the length of the proteins to known conifer diterpene synthases (alignments in Figures A.5 and A.6). A phylogenetic analysis, including conifer mono-, sesqui-, and diterpene synthases, also showed that TpdiTPS1–3 grouped with diterpene synthases (Figure A.7). Among conifer diterpene synthases, TpdiTPS1 and 2 separated at the base from diterpene synthases in the Pinaceae family of conifers to form a clade that grouped together with Pimara-8(14)15-diene synthase from *T. cryptomerioides* (Figure 3.1). This grouping reflects current phylogenetic classifications as both *T. plicata* and *T. cryptomerioides* belong to the Cupressaceae family of conifers. TpdiTPS3 fell in a separate clade together with diterpene synthases from Pinaceae species. Its position at the base of the clade indicates an early divergence from the Pinaceae diterpene synthases.

Assessment of the expression in RNA-seq libraries derived from radial fractions of stems revealed high steady-state levels of transcripts in the heartwood transition zone relative to sapwood alone (Figure 3.2). The average expression of TpdiTPS2 was 10.8 times higher and TpdiTPS3 86.7 times higher in the sap to heartwood transition zone relative to their expression in sapwood only ($p < 0.05$). Although the mean TpdiTPS1 expression was higher in the transition zone relative to the sapwood fractions, it was not statistically significant. TpdiTPS3 transcript levels were also elevated in the bark fraction relative to the sapwood fractions ($p < 0.05$).

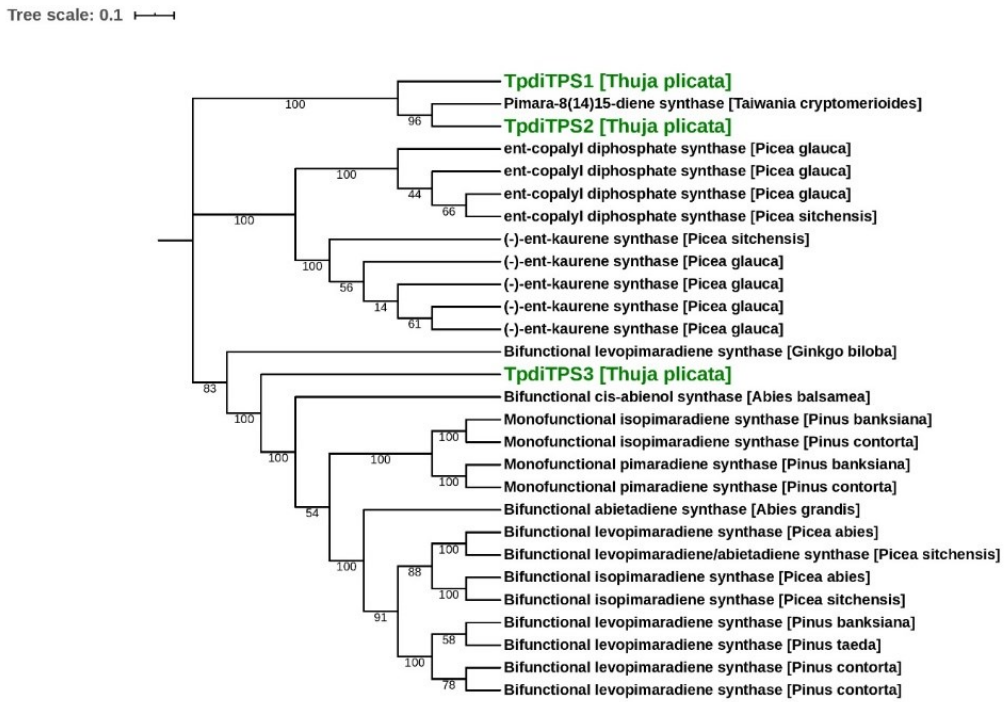


Figure 3.1. Phylogenetic tree of conifer diterpene synthases.

Phylogenetic analysis was performed with 28 full-length diterpene synthases. Proteins identified in this study are highlighted in green. TPS (Terpene Synthase) of the Cupressaceae, except for the TPdiTPS3 cluster apart from TPS of the Pinaceae, indicative of the independent diversification and evolution of specific TPS functions in the Cupressaceae and the Pinaceae. Scale bar indicates 0.1 amino acid substitution per site. The number shows bootstrap confidence values from 100 replicates.

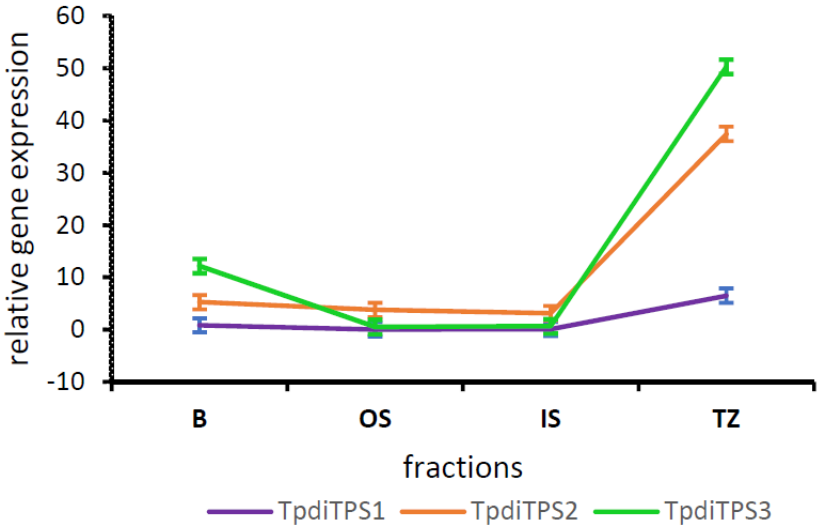


Figure 3.2. TpdiTPS2 and 3 are expressed at elevated levels in the sap to heartwood transition zone.

Expression of *TpdiTPS1, 2,* and *3* in RNA-seq libraries derived from four different fractions of *T. plicata* wood. B—bark, OS—outer sapwood, IS—inner sapwood, TZ—sapwood to heartwood transition zone. Standard error bars were calculated based on two biological replicates. TMM (trimmed mean of M-values) normalized TPM (transcripts per million) values were used to calculate relative expression levels. Statistical differences of expression in different fractions were determined in R (version 4.0.1) by two-way ANOVA and Tukey HSD tests.

To identify functions, we compared the putative catalytic motifs in TpdiTPS1, 2, and 3 (Figure 3.3) with the DXDD motif in the N-terminal domain indicative of a class II terpene synthase and a DDXXD motif in the C-terminal domain indicative of a class I terpene synthase [163,164]. TpdiTPS1 and 2 contained the class I motif DDXXD/E in the C-terminal domain but lacked the class II DXDD motif in the N-terminal domain. Instead, they contained DXDI and DXDV, respectively. These results suggest that TpdiTPS1 and 2 are monofunctional class I enzymes that cannot use GGDP as a substrate but can use CDP as a substrate for diterpene cyclization. The lack of type II activity was confirmed by *in vitro* assays, which showed that recombinant TpdiTPS1 and 2 proteins could not use GGDP as a substrate to synthesize diterpenes (data not shown). In contrast to TpdiTPS1 and 2, TpdiTPS3 matched the class II DXDD consensus motif, but not the class I DDXXD/E motif. This pattern of motifs is typical of monofunctional type II proteins that can use GGDP to synthesize CDP but not carry out the final cyclization to generate diterpenes.

		class II					class I				
protein	class	D	X	D	D		D	D	X	X	D/E
PtLAS	I+II	D	I	D	D		D	D	L	Y	D
PcLAS2	I+II	D	I	D	D		D	D	L	Y	D
TpdiTPS1	I	D	I	D	I		D	D	L	Y	E
TpdiTPS2	I	D	I	D	V		D	D	L	Y	D
PbmIso1	I	D	I	G	V		D	D	L	Y	D
PcmIso1	I	D	I	G	V		D	D	L	Y	D
PgCPS	II	D	I	D	D		R	I	F	F	S
OsCPS*syn*	II	D	I	D	D		S	S	F	H	R
OsCPS2*ent*	II	D	I	D	D		A	S	H	L	R
TpdiTPS3	II	D	I	D	D		S	D	L	Y	E

Figure 3.3. Putative catalytic motifs in TpdiTPS1 and 2 match monofunctional class I proteins and TpdiTPS3 motifs match monofunctional class II proteins.

Amino acid alignments of class I and II motifs of *Thuja plicata* TpdiTPS (1–3) with *Pinus taeda* bifunctional class I/II levopimaradiene synthase (PtLAS, Q50EK2), *Pinus contorta* bifunctional class I/II levopimaradiene/abietadiene synthase2 (PcLAS2, JQ240311), *Pinus banksiana* (PbmIso1, JQ240313), and *Pinus contorta* (PcmIso1, JQ240314) monofunctional class I isopimaradiene synthase1, *Picea glauca* monofunctional class II *ent*-copalyl diphosphate synthase (PgCPS, GU045755), *Oryza sativa* monofunctional class II *syn*-copalyl diphosphate synthase (OsCPS*syn*, AY530101), *Oryza sativa* monofunctional class II *ent*-copalyl diphosphate synthase (OsCPS2*ent*, AY602991). Amino acids that differ from the conserved class II DXDD and class I DDXXD/E are shaded in yellow. Full-length alignments are shown in Figures A.5 and A.6.

3.2.2. Assessment of Potential Mono-Functional Class I Diterpene Synthase Activity in TpdiTPS1 and 2

The lack of an intact DXDD motif in the class II active site of TpdiTPS1 and TpdiTPS2 suggests that these proteins may require CDP rather than GGDP as a substrate. The majority of labdane-related diterpenoid natural products are derived from the *ent*, *syn*, or normal stereoisomers of CDP. As CDP is not commercially available and expensive to chemically synthesize, we used a previously described *in vivo* co-expression system in *E. coli* [285] in which plasmid-based expression of the known proteins result in strains that produce GGDP, *ent*-CDP, *syn*-CDP, and normal-CDP. In these strains, we co-expressed TpdiTPS1, 2, or 3 to test the possibility of bifunctional or monofunctional enzyme activities. The major diterpene product from the co-expression of TpdiTPS1 in a strain that produces normal-CDP was sandaracopimaradiene (Figure 3.4). When expressed in a strain that generates *syn*-CDP as a substrate, TpdiTPS1 produced *syn*-stemod-13(17)-ene as the major product with an unknown diterpene as a minor product (Figure 3.5). Expression of TpdiTPS1 in strains that produce GGDP, or *ent*-CDP did not yield any products (data not shown).

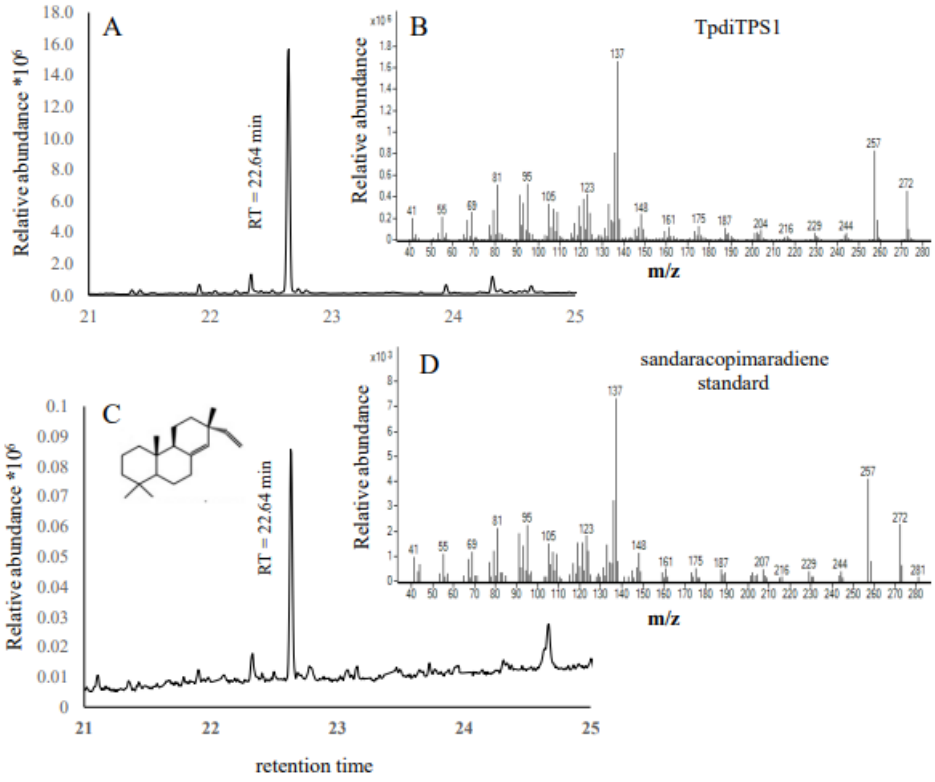


Figure 3.4. GC-MS analysis of sandaracopimaradiene produced by TpdiTPS1 synthase from normal-copalyl diphosphate (CDP).

(**A**) = total ion chromatogram (TIC), (**B**) = mass spectrum of sandaracopimaradiene, (**C, D**) = TIC and mass spectrum of an authentic standard of sandaracopimaradiene produced by *Aspergillus niger* sandaracopimaradiene synthase AnCPS-PS [288].

Expression of TpdiTPS2 with normal-CDP produced levopimaradiene as a major product but did not catalyze any product when incubated with GGDP, *ent*-CDP, and *syn*-CDP (Figure 3.6). The mass spectrum was identical to the levopimaradiene produced by *P. abies* levopimaradiene/abietadiene synthase PaTPS-LAS [289].

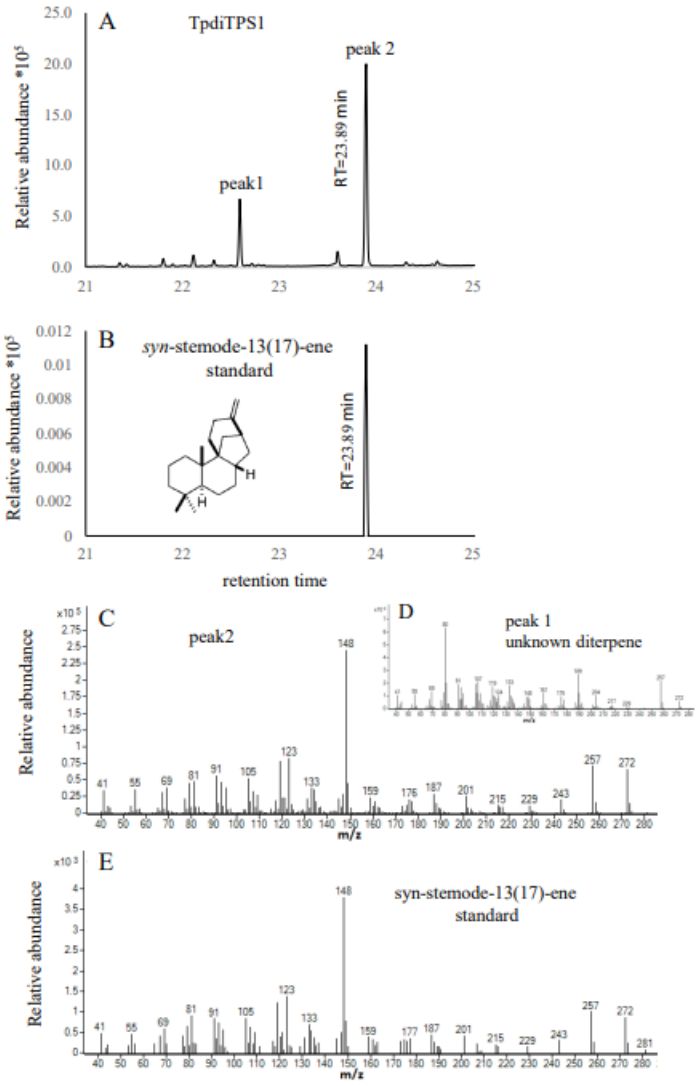


Figure 3.5. GC-MS analysis of syn-stemod-13(17)-ene produced by TpdiTPS1 synthase from syn-copalyl diphosphate (CDP).

(A) = GC-MS total ion chromatogram (TIC) showing unknown diterpene (peak 1) and *syn*-stemod-13(17)-ene (peak 2) produced by TpdiTPS1 from *syn*-CDP.

(B) = TIC of *syn*-stemod-13(17)-ene standard produced by *Oryza sativa syn*-stemod-13(17)-ene synthase OsKSL11 [285]. (C, D) and (E) = mass spectra of peak 2, peak 1, and an authentic *syn*-stemod-13(17)-ene standard produced by OsKSL11 [285], respectively.

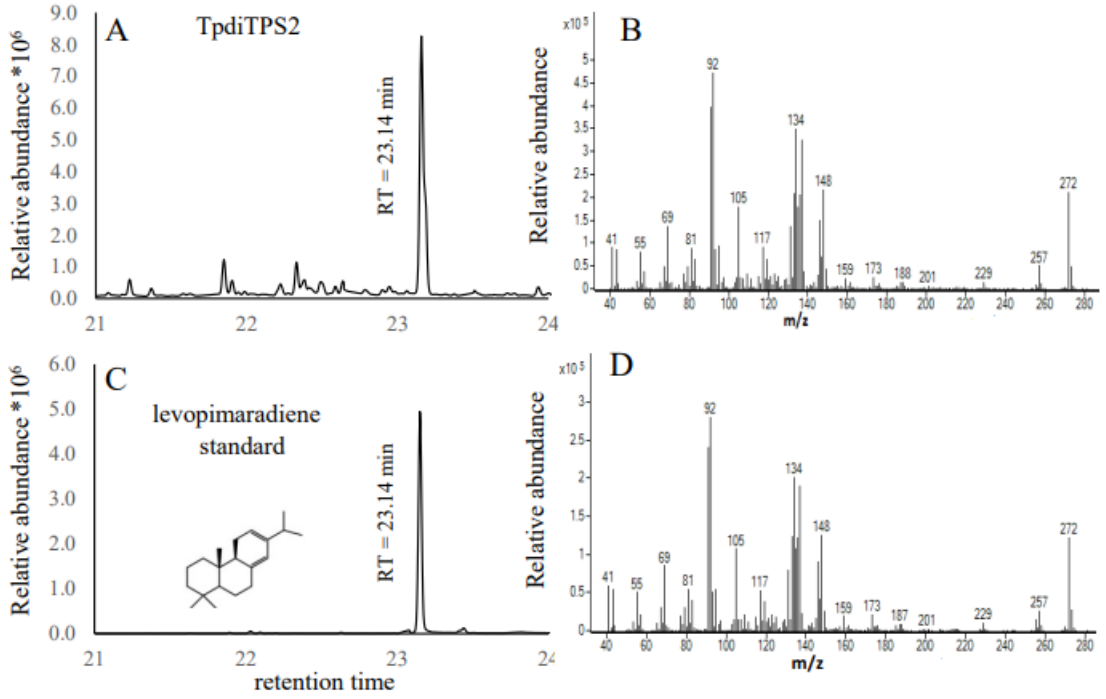


Figure 3.6. GC-MS analysis of levopimaradiene produced by TpdiTPS2 from normal-CDP.

(A) = total ion chromatogram (TIC), (B) = mass spectrum of levopimaradiene, (C, D) = TIC and mass spectrum of an authentic standard of levopimaradiene produced by *Abies grandis* abietadiene synthase AgAS [290].

3.2.3. Monofunctional Class II TpdiTPS3 Synthesize the Intermediate CDP

TpdiTPS3 resulted in copalol when co-expressed with the gene producing GGDP (Figure 3.7), in agreement with being a monofunctional class II enzyme. No product was generated when expressed together with genes that produce *ent*-CDP, *syn*-CDP, and normal-CDP. The TpdiTPS3 product and *syn*-CDP had different retention times and mass spectra. Since *ent* and normal CDP are enantiomers, they are not separated by GC-MS analysis. Therefore, to assess whether the product of TpdiTPS3 is *ent* or normal CDP, we used the *A. thaliana* AtKS (*ent*-kaur-16-ene synthase) enzyme [285] that only reacts with *ent*-CDP to produce *ent*-kaur-16-ene as a means of testing substrate specificity. While expression of an *ent*-CDP synthase with the *ent*-kaur-16-ene synthase resulted in the production of the expected *ent*-kaur-16-ene (Figure 3.8A), the

combination of TpdiTPS3 and *ent*-kaur-16-ene synthase did not (Figure 3.8C), providing evidence by exclusion that TpdiTPS3 is a normal-CDP synthase.

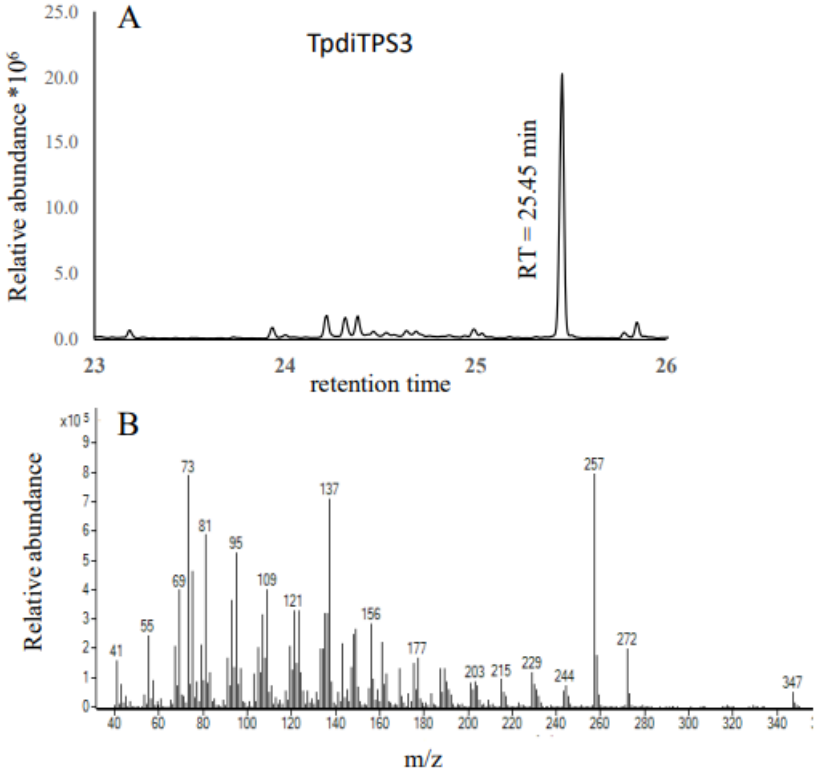


Figure 3.7. GC-MS analysis of copalol produced by TpdiTPS3 from geranylgeranyl diphosphate (GGDP).

(A) = GC-MS analysis of copalol derived from geranylgeranyl diphosphate (GGDP) by TpdiTPS3

(B) = Mass spectrum of the chromatograph peak for TpdiTPS3.

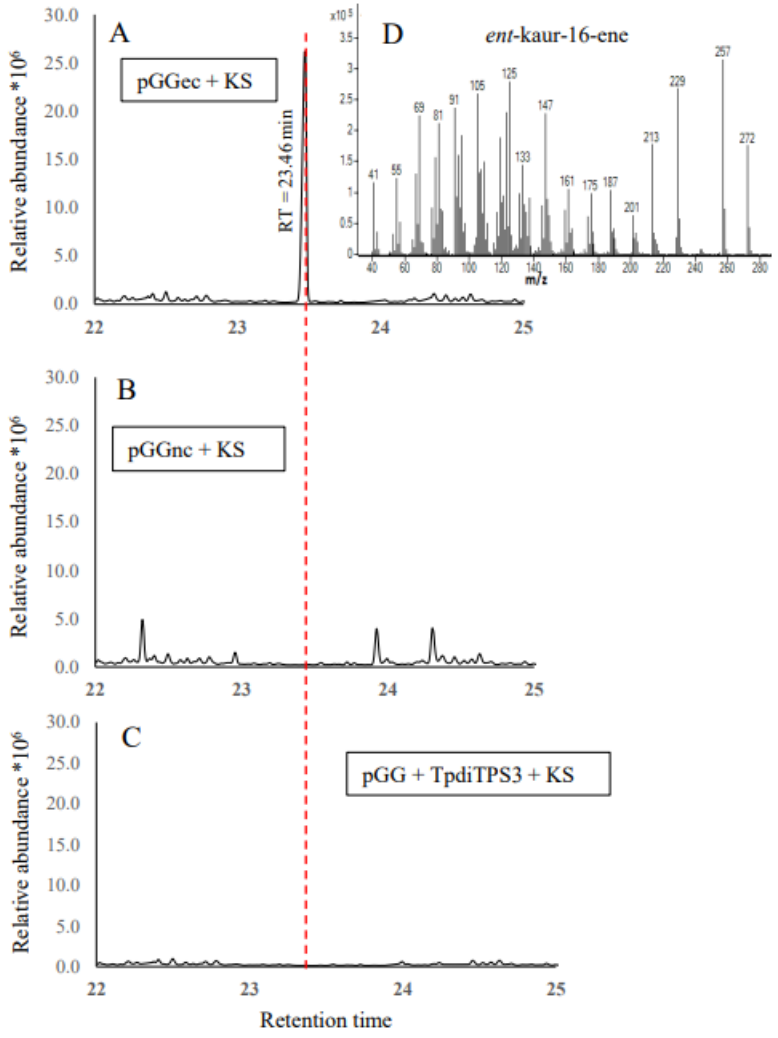


Figure 3.8. Confirmation of enantiospecific enzymatic product of TpdiTPS3.

(A) = Expression of *ent*-CDP synthase with *ent*-specific kaurene synthase results in *ent*-kaur-16-ene biosynthesis. (B) = Expression of *ent*-specific kaurene synthase with normal CDP synthase did not yield any *ent*-kaur-16-ene. (C) = Expression of *ent*-specific kaurene synthase with GGDP synthase and TpdiTPS3 also did not yield *ent*-kaur-16-ene. (D) = Mass spectrum of *ent*-kaur-16-ene.

3.3. Discussion

The genetic basis of terpene biosynthesis during heartwood formation in *T. plicata* is currently unknown. Here we show that three cDNAs derived from mRNAs expressed in the sap to heartwood transition zone of *T. plicata* trees, named TpdiTPS1, 2, and 3, encode diterpene synthases with specific *in vivo* activities. Consistent with the catalytic motif profiles in their peptide sequences, they proved to be mono rather than

bifunctional diterpene synthases. TpdiTPS3 had a DXDD motif in the N-terminal domain known to be required for the protonation-initiated cyclization of GGDP to CDP [169] but deviated from the DDXXD motif in the C-terminal domain, known to be required for the diphosphate ionization-initiated cyclization and rearrangement of CDP to diterpene hydrocarbons [64]. Accordingly, the expression of TpdiTPS3 in the presence of GGDP resulted in the biosynthesis of CDP. A method of exclusion was used to gain insight into what isomer of CDP is produced. The retention time and mass spectrum matched that of normal and *ent*-CDP (indistinguishable from each other) but differed from and excluded *syn*-CDP. Since an *ent*-CDP-specific *ent*-kaurene synthase did not produce *ent*-kaurene when expressed with TpdiTPS3 but did when expressed with an *ent*-CDP synthase, TpdiTPS3 produces normal-CDP; see model, Figure 3.9. Confirmation by an independent method, i.e., chiral column chromatography, is currently beyond our means. However, normal-CDP identity is also consistent with the normal-CDP use of TpdiTPS1 and 2. A class II normal-CDP synthase has been identified before in *Z. mays*, wheat (*Triticum aestivum* L.), *S. miltiorrhiza*, and conifer *T. cryptomerioides* [65–68], which, like *T. plicata*, belongs to the Cupressaceae family of conifers. There are also known bifunctional diterpene synthases that produce normal-CDP as an intermediate [171,289,290,295–297].

The peptide sequences of TpdiTPS1 and 2 lacked an intact N-terminal DXDD motif but contained the C-terminal DDXXD motif required for the rearrangement of CDP to diterpene hydrocarbons, typical of class I diterpene synthases [170,291]. In agreement with their motif profile, both TpdiTPS1 and 2 produced diterpenes when co-expressed with proteins providing CDP. TpdiTPS1 accepted normal and *syn*-CDP as a substrate, but not *ent*-CDP. TpdiTPS1 produced sandaracopimaradiene in the presence of normal-CDP and *syn*-stemod-13(17)-ene in the presence of *syn*-CDP. The use of *syn*-CDP suggests that a *syn*-CDP synthase may be expressed during *T. plicata* heartwood formation. On the other hand, if *syn*-CDP is not available, TpdiTPS1 may be limited to normal-CDP and produce only sandaracopimaradiene during heartwood formation. Normal-sandaracopimaradiene has been found as the sole product from one monofunctional class I diterpene synthases cloned from *O. sativa* [298]. In conifers, monofunctional class I diterpene synthases have been found to produce normal-sandaracopimaradiene only as minor products (*P. banksiana*; 10%, *P. contorta*; 12%) [156], which differs from the >90% normal-sandaracopimaradiene produced by

TpdiTPS1. Bifunctional diterpene synthases with normal-sandaracopimaradiene as minor product have been found in *P. sitchensis* and as a major product in the fungi (*Aspergillus niger* Tiegh.) [146,288,290]. A stemodene synthase has been identified in plants only once before [299], in *O. sativa*. TpdiTPS2 used normal-CDP to produce normal-levopimaradiene. An enzyme with the same specificity was recently identified in another Cupressaceae species, *T. cryptomerioides* [64]. Bifunctional diterpene synthase that produces levopimaradiene as one of several products have been found in ginkgo (*Ginkgo biloba* L.) and other conifers *A. grandis*, *P. taeda* and *P. abies* [289,290,296,297]. However, there is evidence that the four diterpenes identified in *in vitro* assays of the *P. abies* enzyme are epimers from a thermally unstable diterpenol [177,289]. It is possible that we obtained similar artifacts, although we did not see a similar spread of artifact products. Taken together, the substrate and product specificities of TpdiTPS1, 2, and 3 are unusual among plant species and provide the first model framework for diterpene biosynthesis during heartwood formation in *T. plicata* (Figure 3.9).

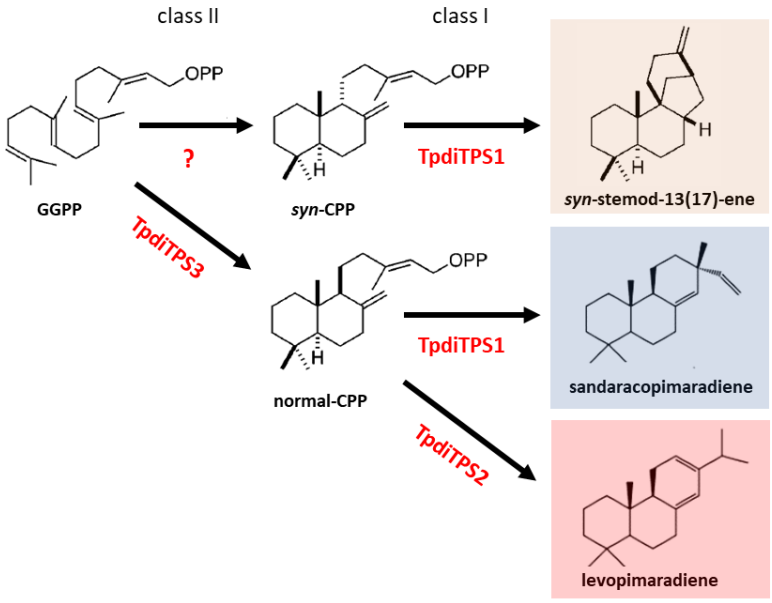


Figure 3.9. Model of the monofunctional diterpene biosynthesis pathway potentially active during heartwood formation in T. plicata trees based on the results in this study.

The question mark indicates that an unidentified *syn*-CDP synthase may also be active to provide an alternative substrate for TpdiTPS1.

To date, no single diterpene or diterpenoid has been described in *T. plicata* tissues [10,61,62,211,212,217,220,300–302], so it is unclear in what modified form we should expect to find the three diterpene products identified in this study. In the closely related species Japanese thuja (*Thuja standishii* Gordon Carrière), extracts from stem bark contain both diterpene alcohols and acids [303], none with the same hydrocarbon skeleton as described in this study. Among other Cupressaceae, the sandarac tree (*Tetraclinis articulata* Vahl Mast.) is well known for its gum, used since antiquity as a natural pictorial and wood varnish [304]. Sandaracopimaric acid and sandaracopimarinol are the two most abundant diterpenoids in sandarac gum [305]. Species in the *Callitris* genus have also been used to produce sandarac gum and have alcohol and acid forms of sanadaracopimaradiene in bark resin [306,307].

Little is known about the toxicity of the diterpenes identified here and their potential natural modified forms. Sandaracopimaric acid, purified from pine cones, inhibits both brown and white rot-causing fungal species in petri-dish experiments. Of nine tested diterpene acids, levopimaric acid generated the strongest growth inhibition of brown and white rot fungal species [308]. Closely related compounds are also found among diterpene phytoalexins in Poaceae species, produced in response to microbial attack [309,310]. Tests demonstrate their toxicity to a range of fungal and bacterial species [311–313]. There is also evidence that related diterpene acids play a role in conifer resistance to insect pests and associated fungi [135].

This chapter aimed to identify genes encoding diterpene synthases that are likely to contribute to heartwood rot resistance in WRC. From our results, we can conclude that there are at least three putative diterpene synthases expressed in the region in which specialized metabolites accumulate to generate heartwood. The enzymes are monofunctional and together provide a potential pathway for the production of levopimaradiene, sandaracopimaradiene, and *syn*-stemod-13(17)-ene from GGDP via the intermediates normal and *syn*-CPP. These are the first terpene synthases identified as active during heartwood formation in WRC and provide a starting point to (1) identify corresponding diterpenoids in heartwood extracts, (2) identify genes that carry out required modifications to the hydrocarbon skeletons, and (3) carry out Single Nucleotide Polymorphism-heartwood rot association studies to test their potential roles in heartwood rot resistance, the single most important economic problem in this species [2,9,259].

Chapter 4. Functional analysis of three putative Fe(II)/2-oxoglutarate dependent dioxygenase genes through heterologous and transient expression assay

4.1. Materials and Methods

4.1.1. Selection of T. plicata dioxygenase (TpDOX) genes

To identify putative TpDOX genes, a *T. plicata* transcriptome prepared from sap to heartwood transition zone (TZ) [263] was searched with a putative “dioxygenase” annotation as a query. Three nucleotide sequences were selected based on the presence of conserved motifs and level of expression (Figure A.8). The BLASTX and BLASTP algorithms were further used to assess their sequence similarity to other known dioxygenase sequences. Multiple alignments of protein sequences were done with MEGA version 11 [314] and the conserved motifs were visualized in Gene Doc version 2.7 [315]. A phylogenetic analysis was done in MEGA version 11, visualized, and modified in iTOL version 5 [316]. DOX genes from other species were retrieved from the National Center for Biotechnology Information (NCBI) GenBank database.

4.1.2. Gene amplification and cloning

RNA extraction and cDNA synthesis from TZ was carried out as described in chapters 2 and 3. The open reading frame (ORF) of three putative dioxygenases were PCR amplified from TZ cDNA with vector overlap primers (Table B.6) according to CloneAmp™ HiFi PCR premix (Takara Bio, Kusatsu, Japan) protocol with annealing temperature of 57°C. Expression vector pET28b (Addgene, Watertown, MA, USA) was digested with *Nco*I and *Xho*I restriction enzymes. The PCR-amplified fragments and the digested vector were purified using GeneJET PCR Purification Kit (Thermo Fisher Scientific). The purified PCR products were cloned into the digested vector by following the protocol of the HD Infusion cloning kit and transformed into Stellar chemically competent cells (Takara Bio, Inc.). Colonies harboring plasmid with insert were identified by colony PCR. The constructs were purified using GeneJET Plasmid Miniprep Kit (Thermo Fisher Scientific) and inserts were confirmed by Sanger sequencing.

4.1.3. Heterologous expression of TpDOX genes in E. coli

Recombinant protein expression and purification were performed as described [146]. For the overexpression of putative TpDOX genes, 200 ng of the ORF-containing plasmid was transformed into 60 µl BL21 (DE3, ThermoFisher, Waltham, MA, USA) chemically competent *E. coli* cells. A plasmid without an insert was transformed as a control. For functional expression, four individual colonies of transformants were inoculated in 5 mL seed culture of Luria-Bertani (LB) medium supplemented with 35 µg/mL Kanamycin, which was cultured overnight at 37°C, at 225 rpm. The following day, 2 mL seed culture was inoculated in 100 mL production culture of Terrific Broth (TB) medium with the same concentration of antibiotic, which was incubated at 37°C at 225 rpm until an OD_{600nm} of 0.8. Cultures were then cooled on ice for 20 min before induction. Protein expression was induced by the addition of 0.2 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG) solution. The induced cultures were incubated overnight at 15°C at 225 rpm. The 100 mL cultured cells were split evenly into two 50 mL falcon tubes and centrifuged at 4°C, at 3270 RCF for 30 min. The supernatant was discarded, and the cell pellets were stored at -80°C until proceeding with protein purification.

4.1.4. Purification of recombinant proteins

Cell pellets (typically 1g/50 mL culture) were resuspended in 2.5 mL of freshly prepared ice-cold lysis buffer. The resuspended pellet was placed on ice for 30 min with occasional stirring with a glass rod. The suspension was transferred into pre-cooled 15 mL falcon tubes and lysed by sonication on ice for total of 2 min (30% strength, 3x15 s, 30 s off between, followed by second 3x15 s, 30 s off between). The cleared lysate was transferred into pre-cooled 2 mL microcentrifuge tubes and centrifuged for 30 min, at maximum RCF, at 4°C. The supernatant was loaded onto a His Spin Trap Ni-affinity column (GE Healthcare, Chicago, IL, USA). The purification steps were according to the manufacturer's protocol. The recombinant proteins were eluted with 300 µl elution buffer and the protein concentration was determined on Nanodrop (ThermoFisher, Waltham, MA, USA). Purified proteins were desalted using PD Mini-Trap G-25 desalting columns (GE Healthcare) following the manufacturer's protocol. Protein concentration was further determined on a nanodrop instrument (ThermoFisher, Waltham, MA, USA). The purified proteins were immediately used for enzyme assay and the rest of the proteins were split into aliquots of 50 µl and flash-frozen in liquid N_2 for storage at -80°C.

4.1.5. SDS-PAGE analysis

Purified protein samples were examined for overexpression on 4-15% Mini-PROTEAN TGX (Tris-Glycine eXtended shelf life) stain-free gel (Bio-Rad Laboratories Inc., Montreal, QC, Canada). At first, 50 µl β-mercaptoethanol was added fresh in 950 µl sample buffer from which 7 µl was mixed with 10 µl protein (~40 µg). This mixture was heated to 95°C for 5 min and placed on ice until required. The total 17 µl sample was loaded onto the precast gel alongside 10 µl of PageRuler Unstained Protein Ladder (Thermo Scientific) and run on Mini-PROTEAN II electrophoresis cell with Bio-Rad power supply (Model 1420B/150). The gel was run at 300 V until the smallest protein ladder band reached the bottom of the gel (~1.5 h). The protein was visualized with Bio-Rad's stain free instrument which includes a ChemiDoc MP imager. The imager was connected to a separate computer and run by Image Lab software.

4.1.6. In Vitro enzyme assay

Stock solutions of 2-oxoglutaric acid (2OG), sodium ascorbate and ammonium iron (II) sulfate hexahydrate were prepared fresh in sterile dH_2O. These solutions were not stored for more than 3 h. The stock solution of cymorcin was prepared in DMSO or methanol. Enzyme assays were performed in 100 µl reaction containing 50 mM of Tris-HCl or TES buffer, pH 7.5 (25 µl), 5 mM 2-oxoglutaric acid (5 µl), 0.1 mM ammonium iron (II) sulfate hexahydrate (1 µl), 8 mM sodium ascorbate (8 µl), 2.5 mM cymorcin (1 µl), 10 µl (30-40 µg) enzyme and 50 µl sterile dH_2O. Assays were incubated at 30°C for 2 h. During incubation, vials were left open to get adequate atmospheric O_2. The reactions were acidified to pH 2 by the addition of 20 µl of 3M HCl. After acidification, the samples were extracted with 400 µl ethyl acetate (EtOAc) and hexane (3:1) by vortexing for 30 s. The samples were centrifuged at 1000 g for 30 min at 4°C to separate the phases. The top 300 µl organic fractions were transferred in a separate vial and evaporated completely under the flow of nitrogen and re-dissolved in 100 µl EtOAc. The samples were derivatized with 15 µl N,O-Bis(trimethylsilyl)trifluoroacetamide (BSTFA) and incubated for 1 h at 70°C prior to Gas Chromatography-Mass Spectrometry (GC-MS) analyses. The above *in vitro* assay was also carried out for substrate 4(8)-p-menthene-1,2-diol.

4.1.7. In Vivo enzyme assay

For *in vivo* assay, pET28b constructs were transformed into BL21 (DE3, ThermoFisher, Waltham, MA, USA) *E. coli* cells as described above. An empty plasmid without an insert was transformed as a control. Four individual colonies of transformants were inoculated in 5 mL Luria-Bertani medium supplemented with 35 µg/mL Kanamycin and were grown overnight at 37°C, 225 rpm. A fresh overnight culture (1 mL) was inoculated in 50 mL of 2TY medium with the same concentration of antibiotic. The cultures were incubated at 37°C, 225 rpm until an OD_{600nm} of 0.8 (typically 2-3 h). Before induction, cultures were cooled on ice for 20 min. Enzyme assays were performed after induction of gene expression with 0.2 mM (50µl) Isopropyl β-D-1-thiogalactopyranoside (IPTG) was completed. 10 mM ammonium iron (II) sulfate hexahydrate (125 µl), 100 mM sodium ascorbate (125 µl), 100 mM 2-oxoglutaric acid (125 µl) and 200 µl cymorcin (stock 10 mg/mL in DMSO or methanol) were added. The cultures were incubated for 48 h at 15°C, 225 rpm. After centrifugation at 4°C, 1400 RCF for 25 min, both supernatant and cell pellets were separated for metabolite extraction. To extract metabolites from the cell pellet, 25 mL of acetone was added into the pellet and stirred for 3 h before gravity filtration to clarify the solvent. The filtrate was concentrated under vacuum until it was reduced to approximately 4 mL. It was redissolved in 25 mL of sterile dH_2O and acidified to pH 2 with 200 µl of 3M HCl. This solution was extracted with 25 mL of EtOAc and kept at 4°C for overnight. Approximately 16 mL of the organic phase was evaporated completely under N_2 stream, re-dissolved in 100 µl EtOAc and derivatized as described above. The supernatant, 50 mL, was clarified by vacuum filtration, acidified to pH 2 with 2.4 mL of 3 M HCl from which 25 mL of solution was extracted with an equal volume of EtOAc. To separate the phases, samples were kept at 4°C for overnight. Approximately 16 mL of the organic phase was evaporated completely under N_2 stream, re-dissolved in 100 µl EtOAc and derivatized as described above. The above *in vivo* assay was also carried out for substrate 4(8)-p-menthene-1,2-diol.

4.1.8. GC-MS analysis

GC-MS analysis was performed on an Agilent 7890B gas chromatograph (GC) and 5977A mass spectrometer (MS) detector (Agilent Technologies, Santa Clara, USA). Mass spectra were recorded at 70 eV. One microliter of the extract was injected, and compounds were separated on a DB5 column that was 30 m long and 0.25 mm in

diameter and had a film thickness of 0.25 μm. Helium was used as a carrier gas with a flow rate of 1 mL/min. The inlet temperature was 250°C, splitless mode. The GC temperature program started at 50°C, hold 5 min, and increased with a ramp of 10°C/min up to 280°C, 10 min hold. The products were identified by matching retention times and spectra to those of authentic standards. Authentic cymorcin and 4(8)-p-menthene-1,2-diol were purchased from Toronto Research Chemicals (TRC), Toronto, Ontario and β-thujaplicin was from MilliporeSigma, Burlington, USA.

4.1.9. Callus culture

Seeds of *C. lusitanica* and Kashmir cypress (*Cupressus cashmeriana* Royle ex Carrière) were obtained from Sheffield's Seed Co., Inc. NY, USA. The seeds were surface sterilized in a solution of 20% commercial bleach (10 mL), sterile dH_2O (40 mL) and Tween 20 (100 µl) with constant shaking for 30 min. The solution was poured off and seeds were washed three times with sterile dH_2O. Seeds were dried on sterile filter paper and plated on solid ½ MS medium. The plates were incubated at 4°C for 30 days and thereafter incubated at ambient temperature under light for germination. Approximately 24 days later, young seedlings were cut into small pieces and transferred onto solid callus induction medium (CIM; B5 salts and vitamins, pH 5.5-5.8, 7 g/L phytoagar, 0.01 µM BAP, 10 µM NAA) supplemented with 0.1 mg/mL Timentin. Hormones and antibiotics were added to medium just before pouring the plates. For callus induction, plates were wrapped with foil paper and incubated at 25°C in the dark. After one month, the emerging callus was chopped into small pieces and transferred onto fresh CIM supplemented with 0.1 mg/mL Timentin. The callus was subcultured every 2-4 weeks and plates kept at 25°C in the dark.

4.1.10. Transformation of TpDOXs in Agrobacterium strain

Three putative dioxygenases were PCR amplified from their corresponding pET28b constructs by using CloneAmp™ HiFi PCR premix (Takara Bio, Kusatsu, Japan) kit using 54°C annealing temperature. The primers had 5' ends that overlap with pWBVec8-Ubi:USER:NOS ends after digestion with *PacI* enzyme (Table B.7). PCR amplified fragments and the digested plasmid were purified using GeneJET PCR Purification Kit (Thermo Fisher Scientific). The purified PCR products were cloned into the plasmid using HD Infusion cloning enzyme and transformed into Stellar chemically

competent cells (Takara Bio, Inc.). An empty plasmid was also transformed as a control. Colonies harbouring plasmid was confirmed through colony PCR and Sanger sequencing. Purified plasmids were electroporated into the *Agrobacterium* strain GV3101. The cells were plated in LB medium containing 100 µg/mL spectinomycin and incubated at 28°C for 3-4 days. Afterwards, colonies were streaked onto solid LB medium containing 100 µg/mL spectinomycin and 40 µg/mL gentamycin and grown for an additional 2-3 days under the same condition.

4.1.11. Agrobacterium-mediated transient expression of TpDOXs in callus

A single colony of *Agrobacterium* was inoculated in 4 mL LB medium containing 100 µg/mL spectinomycin and 40 µg/mL gentamycin. The cultures were grown overnight at 28°C at 240 rpm. 1.5 mL of overnight culture was diluted in 25 mL LB medium containing the above antibiotic concentration and grown at 28°C at 240 rpm until OD_{600} reached to 0.6. The cells were harvested by centrifugation at 2,599 RCF for 10 min at 20°C. The supernatant was discarded, and the cell pellet was resuspended in ~22 mL of liquid CIM and 150 µM acetosyringone. The final OD_{600} was adjusted to 0.6. For co-cultivation, 300 µl of *Agrobacterium* suspension was mixed with ~0.45 g of *C.lusitanica* and *C. cashmeriana* callus. In no *Agrobacterium* control, calli were only mixed with liquid CIM and 150 µM acetosyringone without any *Agrobacterium* culture. The calli were left immersed for 30 min in the dark at 28°C. Excess bacteria were blotted dry on sterile filter paper and transferred to solid CIM and 150 µM acetosyringone (Co-cultivation media). The plates were sealed with parafilm and incubated in dark at 28°C for 48 h. After co-cultivation, calli were wrapped in foil paper and stored at -80°C for RNA extraction.

4.1.12. Confirmation of overexpression via RTqPCR

4.1.12.1. RNA extraction

Calli were ground with a mortar and pestle in the presence of liquid nitrogen. Approximately 0.24 g of ground callus was transferred in 2 mL centrifuge tube and extracted with 1 mL pre-warmed (65°C) extraction buffer (2% CTAB; 200 mM Tris, pH

8.0; 50 mM EDTA, pH 8.0; 2.5 M NaCl; 0.5% activated charcoal; 1.5% polyvinylpolypyrrolidone and 1.5% β-mercaptoethanol) by incubating at 65°C for 10 min. The homogenate was extracted three times with an equal volume of chloroform and centrifuged at 10,625 RCF for 5 min, at 4°C. The aqueous phase was transferred into a new tube and ¼ volume of 10 M LiCl was added followed by incubation on ice for overnight at 4°C. The tubes were centrifuged at 14,462 RCF for 30 min at 4°C. The supernatant was discarded, and the cell pellet resolubilized in 0.5 mL SSTE (1 M NaCl; 0.5% SDS; 10 mM Tris-HCl, pH 8.0; and 1 mM EDTA, pH 8.0) buffer followed by extraction with 0.5 mL chloroform. Phases were separated at 14,462 RCF for 10 min, at 4°C and the upper phase was collected in a new sterile tube. For the precipitation of nucleic acid, two volumes of 100% ethanol at room temperature were added, and the solution was mixed and incubated at −20°C for 3-4 h. The tubes were centrifuged at 14,462 RCF for 30 min at 4°C. The supernatant was discarded, and the cell pellet was washed with 1 mL of 75% cold ethanol. The tubes were centrifuged at 14,462 RCF, for 5 min, at 4°C, supernatant discarded, followed by additional centrifugation for 2 min. Residual ethanol was pipetted off, the pellet was air dried, and resuspended in 20 µl of diethylpyrocarbonate (DEPC)-treated water. RNA integrity was analyzed by bleach gel electrophoresis [262].

4.1.12.2. Real-Time quantitative PCR

One µg of RNA was DNase treated with RapidOut DNA Removal Kit (Thermo Fisher Scientific). cDNA was synthesized from RNA with OneScript Plus cDNA synthesis kit (ABM). Primers for qPCR were designed using Eurofins Genomics qPCR primer design tool (Table B.8). The qPCR reaction was set up according to the manufacturer's protocol of BrightGreen 2X qPCR Mastermix-No Dye (ABM) in a volume of 20 µL. qPCR parameters were set to 3 step amplification: preincubation at 95°C for 600 s followed by 45 cycles of denaturation at 95°C for 15 s, annealing 56°C for 15 s and extension 72°C for 15 s. Target genes were amplified along with reference *C. lusitanica* 18S rRNA gene in a qPCR machine (Light Cycler 96 system, Roche). Expression levels were measured using the $2^{-\Delta\Delta Ct}$ method.

4.1.13. Extraction of metabolites from calli

The extraction was done as described before [61]. One mL of anhydrous ethanol containing 50 ppm heptadecane internal standard was added to 0.1 g ground powder of calli and mixed well by vortexing. The extraction was carried out in an ultrasonic bath for 1 h during which time the bath temperature rose from approximately 20°C to a final temperature of 40°C. The tubes were centrifuged at 295 RCF for 10 min at 4°C. The supernatant was pipetted off in a new tube without touching pellet and filtered it through quartz wool stuffed into the tapered end of Pasteur pipets. An extra 0.5 mL of ethanol was used to wash the residual calli. The eluents were collected in 2 mL GC-MS vials and sent to the Pacific forestry centre, Victoria for GC-MS analysis and β-thujaplicin quantification. The GC-MS programme was as following: Zebron ZB-5MSplus column (30 m X 0.25 mm, film thickness 0.25 μm). The GC temperature program started at 60°C, hold 5 min, and increased with a ramp of 12°C /min up to 300°C, 20 min hold.

4.1.14. Transformation of callus cells with TpDOX-expressing T-DNA

A few microliters of glycerol stock of *Agrobacterium* containing binary plasmid were inoculated in 4 mL LB medium containing 100 µg/mL spectinomycin and 40 µg/mL gentamycin and grown for overnight at 28°C at 240 rpm. 1.5 mL of overnight culture was diluted in 25 mL LB containing above antibiotics and grown at 28°C at 240 rpm until OD_{600} reached to 0.5. The cultures were centrifuged at 2,599 RCF, for 10 min at 20°C. The supernatant was discarded, and the cell pellet was resuspended in ~ 22 mL of liquid CIM and 150 µM acetosyringone. The final OD_{600} was adjusted to 0.5. Afterwards, 1 g of calli were chopped with a scalpel in which 100 µl of *Agrobacterium* suspension was added, mixed, and incubated at 23°C for 30 min. The calli were drained of excess liquid on sterile filter paper and transferred as little clumps on top of sterile filter paper on solid CIM and 150 µM acetosyringone (Co-cultivation media). The plates were sealed with parafilm and incubated in dark at 28°C. After 24 h of incubation, the filter paper was changed with a fresh sterile filter paper and further incubated at the same condition for another 24 h. For stable transformation, calli were thinly spread on sterile filter paper on solid CIM and timentin (400 µg/mL) and incubated at 23°C for 2-3 days to kill bacteria. In cases of *Agrobacterium* overgrowth, calli were washed first with 3-4 mL of sterile water

and then 3-4 mL with water containing 500 µg/mL timentin solution, followed by drying on filter paper and plating on B5 medium. After 2-3 days, transferred the calli without filter paper on solid CIM with timentin (400 µg/mL), and hygromycin (3 µg/mL) and incubated at 23°C for 6 days (Selection Media). After 6 days, the calli were subcultured on fresh solid CIM with cefotaxime (300 µg/mL), and hygromycin (3 µg/mL) and subcultured every 2 weeks.

4.2. Results

4.2.1. Identification of Thuja plicata Fe(II) and 2OG Dependent dioxygenase genes

Three contiguous cDNA sequences encoding putative dioxygenases were retrieved from a *T. plicata* stem transcriptome based on expression level in the transition zone as well as their conserved motifs [263]. These sequences were temporarily named as *Thuja plicata* dioxygenase 1 (TpDOX1), *Thuja plicata* dioxygenase 2 (TpDOX2) and *Thuja plicata* dioxygenase 3 (TpDOX3). The contiguous cDNA sequences of TpDOX1, 2, and 3 were 1661bp, 1132bp, and 1394bp in length, along with open reading frames encoding 341, 350 and 350 amino acid-residue proteins, respectively. The molecular masses are predicted to be 39.19, 40.56 and 40.52 kDa. The multiple protein sequence alignment of TpDOX1, 2 and 3 with other known dioxygenases identified a maximum 30% similarity with *Talaromyces stipitatus* Fe(II)/2-oxoglutarate-dependent dioxygenase (TropC), *Arabidopsis thaliana* gibberellin 2-β-dioxygenase (AtGA2ox1), *Ginkgo biloba* flavonol synthase (GbFLS) and maximum of 36% similarity with *Salvia miltiorrhiza* Fe(II)/2-oxoglutarate-dependent dioxygenase (Sm2OGD25) (Figure A.9). A phylogenetic analysis including other known dioxygenases involved in primary and secondary metabolite biosynthesis revealed that TpDOX1, 2 and 3 separated from dioxygenases involved in primary metabolite biosynthesis (i.e., jasmonic acid, gibberellin, ethylene, IAA). Among dioxygenases involved in secondary metabolite biosynthesis (i.e., tropolone, flavonoid, diterpenoid, glucosinolate), TpDOX1, 2 and 3 clustered specifically with fungal tropolone biosynthesis gene TropC (Figure 4.1).

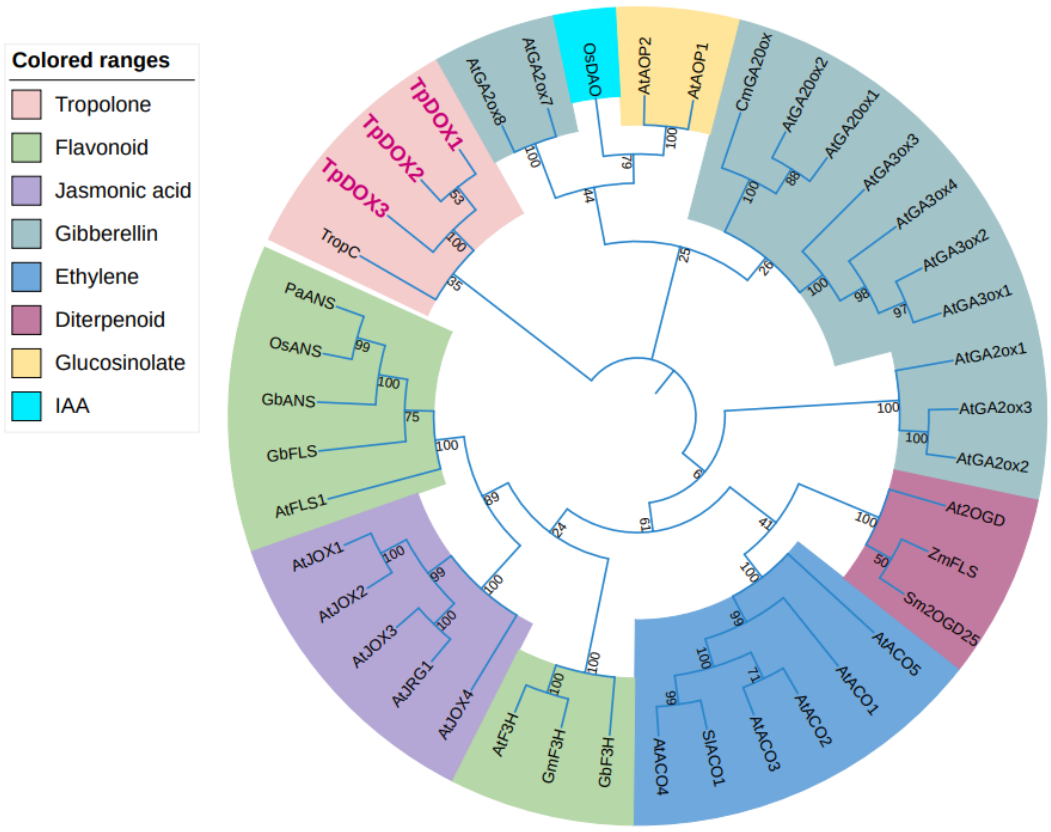


Figure 4.1. Phylogenetic tree of Fe(II) and 2OG dependent dioxygenases.

Phylogenetic analysis was performed with 41 full-length Fe(II) and 2OG dependent dioxygenases. Proteins identified in this study are highlighted in red. TpDOX1-3 cluster with TropC, involved in tropolone biosynthesis, and apart from other functional categories, indicative of the independent diversification and evolution of specific dioxygenase functions. Numbers indicate bootstrap support of 100 repetitions. The full abbreviation of the protein sequences and corresponding accession numbers are listed in Table B.9.

4.2.2. Predicted proteins of TpDOX1,2 & 3 exhibit the conserved motifs of Fe(II)/2OG dependent dioxygenases

The predicted open reading frames (ORFs) of the TpDOX1, 2 & 3 were aligned with other known Fe(II)/2OG dependent dioxygenase ORFs. As shown in figure 4.2, the predicted proteins contain the signature motif HXD/EXnH, which is crucial for Fe(II) binding, and the RXS motif, required for 2OG binding. This analysis suggests that the predicted TpDOX1, 2 & 3 proteins carry the conserved motifs expected for catalytic functions of Fe(II)/2OG dependent dioxygenases.

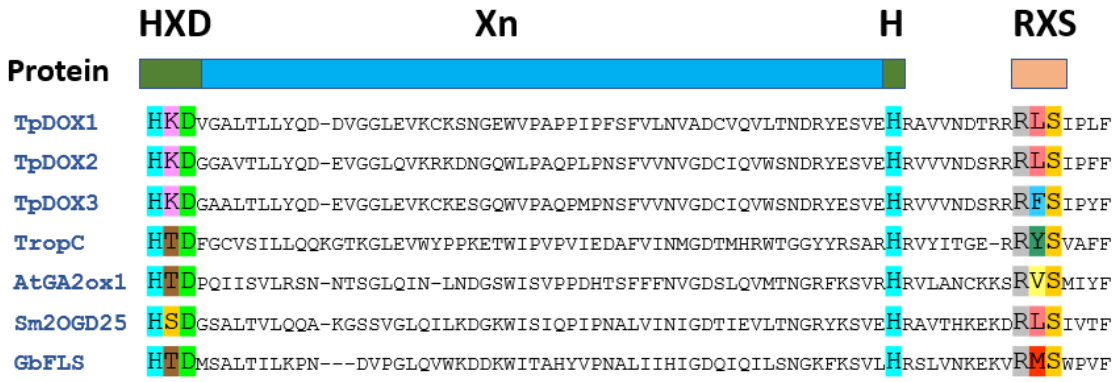


Figure 4.2. Putative catalytic motifs of TpDOX1, TpDOX2 and TpDOX3.

Amino acid alignments of TpDOX1, TpDOX2 and TpDOX3 with *Talaromyces stipitatus* Fe(II)/2-oxoglutarate-dependent dioxygenase (TropC), *Arabidopsis thaliana* gibberellin 2-β-dioxygenase (AtGA2ox1), *Salvia miltiorrhiza* Fe(II)/2-oxoglutarate-dependent dioxygenase (Sm2OGD25), and *Ginkgo biloba* flavonol synthase (GbFLS). Full-length alignments are shown in Figure A.9.

4.2.3. Expression level of TpDOX1,2&3 in the sap to heartwood transition zone

Differential expression analyses showed that transcript levels of TpDOX1 were significantly (p value <0.001) elevated in the transition zone (TZ), relative to other stem fractions (Figure 4.3A). TpDOX1 transcript level was 185 times higher in TZ relative to the average levels in bark (B), outer sapwood (OS), and inner sapwood (IS). TpDOX1 transcript levels are 341, 358 and 95 times higher in TZ compared to B, OS, and IS. The transcript levels of TpDOX2 were about the same in all four fractions and only slightly higher in TZ with no statistically significant differences (Figure 4.3B). TpDOX3 was only 4 times higher in TZ relative to B and its expression is like OS and IS and their differences are also not significant (Figure 4.3C).

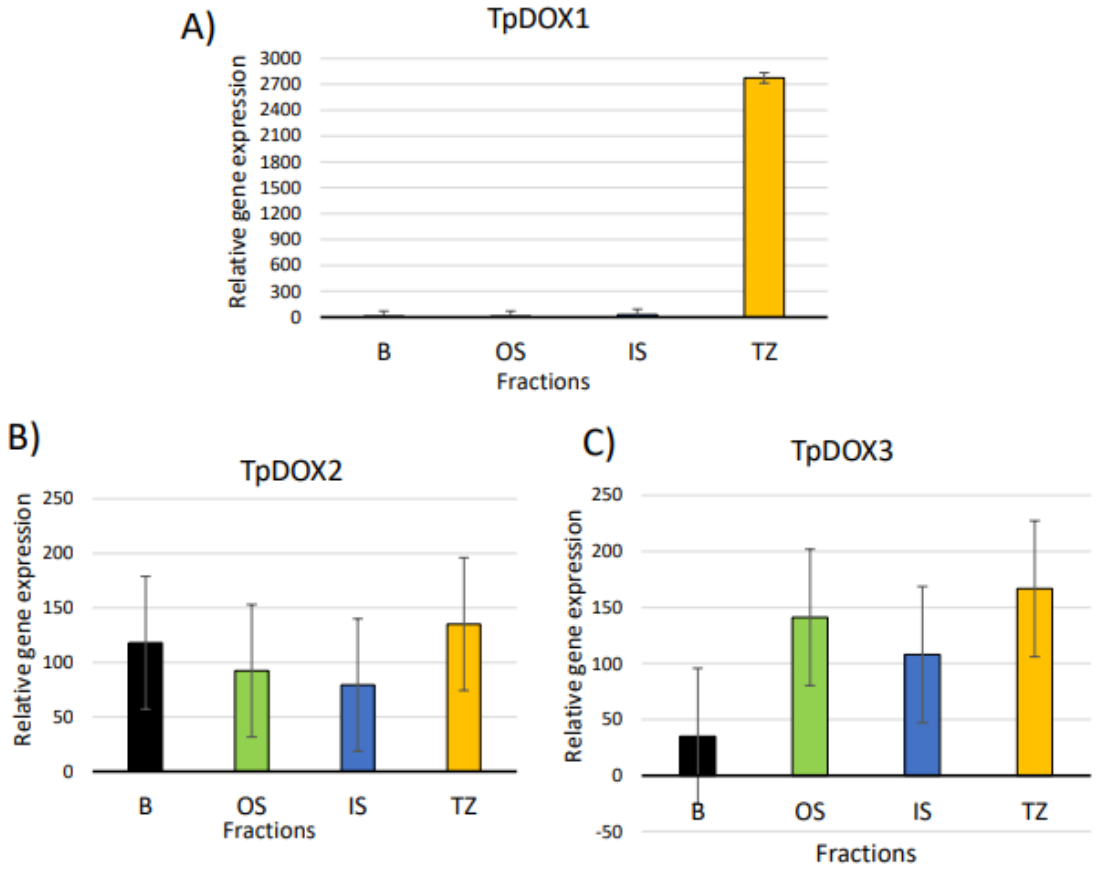


Figure 4.3. Expression levels of TpDOX1, TpDOX2 and TpDOX3 in the sap to heartwood transition zone.

Expression of *TpDOX1*, *TpDOX2* and *TpDOX3* in RNA-seq libraries derived from four different fractions of *T. plicata* wood. B-bark, OS-outer sapwood, IS-inner sapwood, TZ-sapwood to heartwood transition zone. TMM (trimmed mean of M-values) normalized TPM (transcripts per million) values were used to calculate relative expression levels. Statistical differences in different fractions were determined in JMP 15.0 by two-way ANOVA and Tukey HSD tests. Standard error bars were calculated based on two biological replicates.

4.2.4. IPTG induction resulted in overexpression of dioxygenase genes in E. coli

After purification, proteins were run on SDS-PAGE gel. The analysis of the protein content of *E. coli* transformed with pET28b-TpDOX1, pET28b-TpDOX2 and

pET28b-TpDOX3 revealed the presence of a strong band, consistent in size with the calculated mass of His6-TpDOX1 (40.37 kDa), His6-TpDOX2 (41.74 kDa) and His6-TpDOX3 (41.69 kDa). The putative bands were observed in samples induced with IPTG and no strong band of equivalent size was observed in empty vector sample (Figure 4.4).

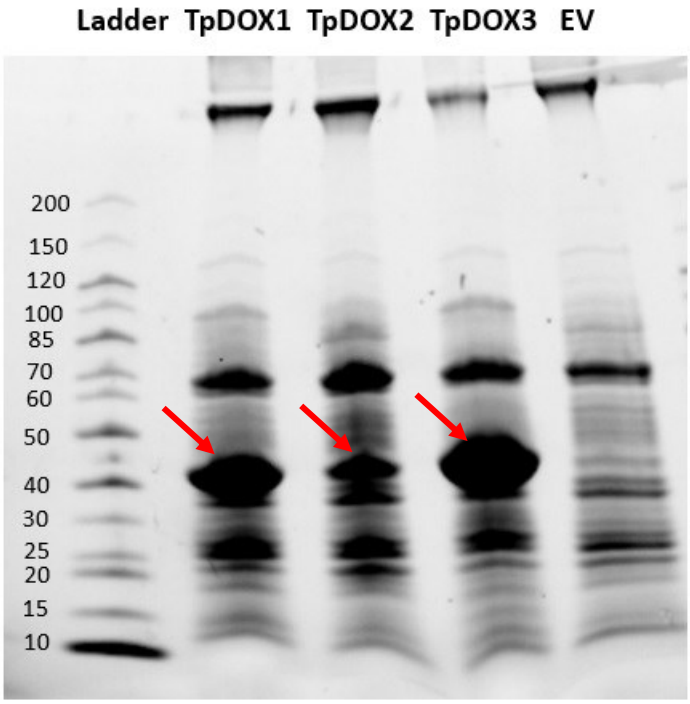


Figure 4.4. SDS-PAGE analysis of the proteins overexpressed in E. coli pET28b system by IPTG induction.

From left to right: Protein molecular weight marker (Ladder, KDa), Red arrow shows the His-tag enriched protein from expression of TpDOX1 (40.37 KDa), TpDOX2 (41.74 KDa), TpDOX3 (41.69 KDa) and Empty vector (EV).

4.2.5. No activity identified from in vitro and in vivo assays of enzymes

To determine the function of TpDOX1, 2 and 3, cDNAs were cloned into an expression plasmid, proteins were produced in *E. coli* cells, and Histidine-tagged recombinant proteins were purified. Enzyme assays using the TpDOX1, 2 and 3 proteins, putative substrate cymorcin, cofactor Fe(II) and cosubstrate 2-oxoglutaric acid yielded no product (Figure 4.5) by GC-MS analysis. An alternative *in vivo* assay in *E. coli*

adding cymorcin, cofactor and cosubstrate also resulted in no product from the dioxygenases. Only derivatized cymorcin was detected in all samples and no additional peak was observed in the samples, indicating that cymorcin is not the right substrate for TpDOX1, 2 or 3 enzymes. Extracts prepared from the same *E. coli* strain transformed with a plasmid lacking cDNA insert also did not yield any detectable compound.

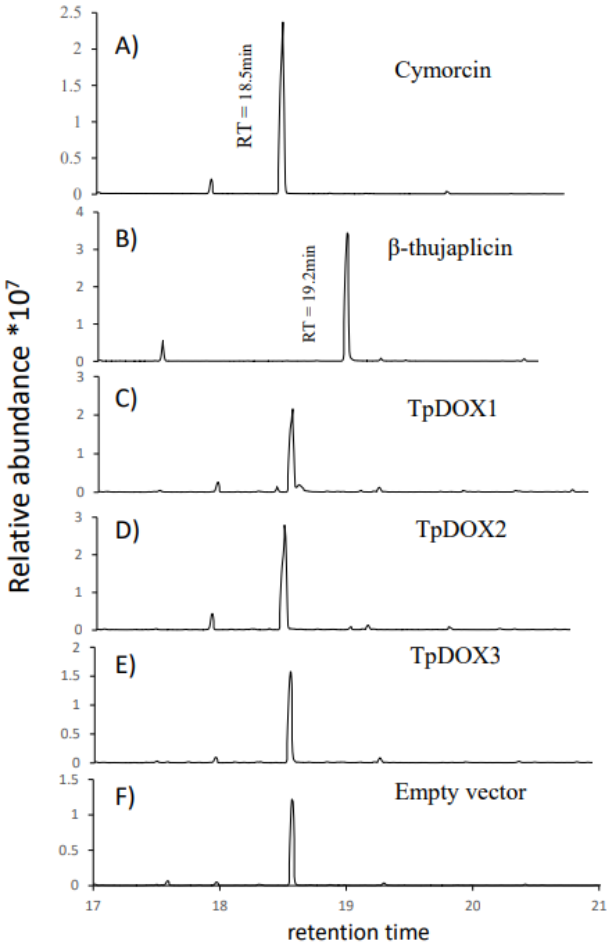


Figure 4.5. TpDOX1, 2 and 3 incubations with cymorcin yielded no product.

Gas chromatography retention times and relative abundance of (A) = standard cymorcin. (B) = standard β-thujaplicin. (C) = Incubation of TpDOX1 with cymorcin. (D) = Incubation of TpDOX2 with cymorcin. (E) = Incubation of TpDOX3 with cymorcin. (F) = Incubation of Empty vector with cymorcin. The mass spectra of derivatized authentic cymorcin and β-thujaplicin appear in Figure A.10.

4.2.6. A. tumefaciens-mediated transient expression in C.lusitanica & C. cashmeriana calli

Neither tissue culture nor *Agrobacterium*-mediated transformation have been established for *T. plicata* and our own attempts have been unsuccessful (data not shown). Since it is known that calli can be induced and maintained from related species, we generated calli from *C. lusitanica* and *C. cashmeriana* seedlings. Attempts at generating *Agrobacterium* T-DNA transformed calli based on hygromycin selection were unsuccessful (data not shown). For TpDOXs overexpression, we therefore aimed at infection and transient expression in a subpopulation of cells similar to Agroinfiltration in *Nicotiana benthamiana* leaves [317]. To that effect, *Agrobacterium* strain GV3101 harbouring TpDOX1, 2, 3 and TpTS1 overexpression plasmids (along with empty vector, EV, and no *Agrobacterium* as negative controls) were used to infect *C. lusitanica*, and *C. cashmeriana* calli. The RNA of the transformed calli was extracted and analyzed by quantitative PCR. In *C. lusitanica* cells, over expression of TpDOX1 or TpTS1 was not detected (not shown). TpDOX2 transcript level was 317-fold higher and ToDOX3 transcript level was 10-fold higher in cells exposed to overexpression plasmid compared to EV sample (Figure 4.6). In *C. cashmeriana*, compared to EV, TpDOX1,2,3 & TpTS1 were 43, 16, 5, and 13-fold higher in cells infected with corresponding overexpression T-DNA, relative to empty-vector transformed cells (Figure 4.7).

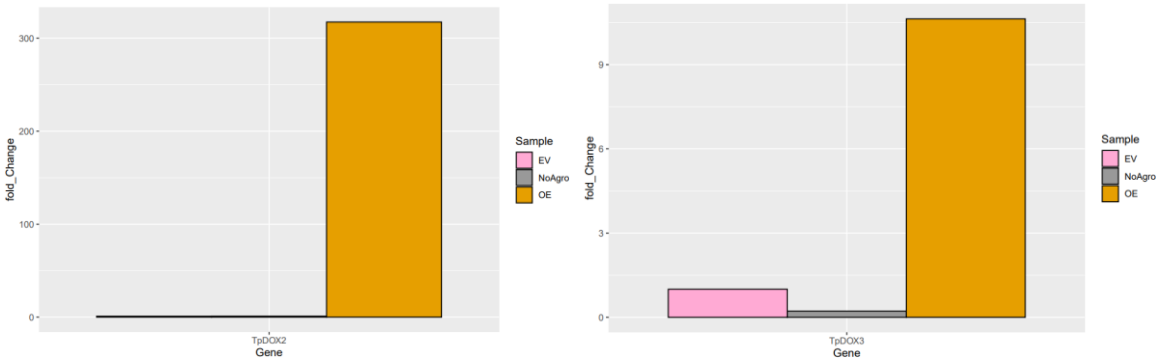


Figure 4.6. Overexpression of TpDOX2 (left) and TpDOX3 (right) in C.lusitanica cells.

RT-qPCR analysis, EV= Empty Vector, NoAgro= No *Agrobacterium*, OE= overexpressed.

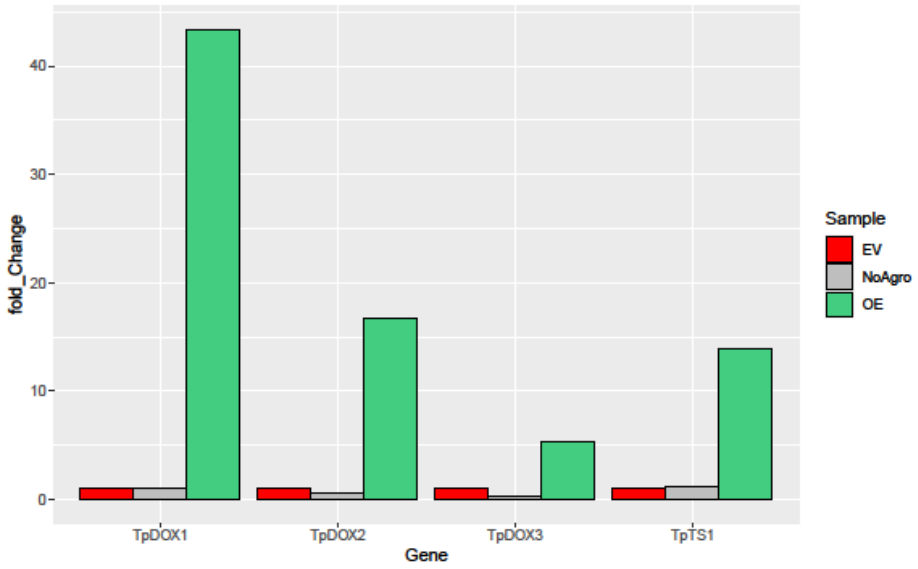


Figure 4.7. **Overexpression of TpDOX 1,2,3 and TpS1 in *C. cashmeriana* cells.**

RT-PCR, EV= Empty Vector, NoAgro= No Agrobacterium, OE= overexpressed.

4.2.7. Extraction and determination of β-thujaplicin in overexpressed Calli

The overexpressed and control samples were grounded in liquid nitrogen and extracted with ethanol. The supernatant was analyzed by GC-MS to quantify β-thujaplicin. The result has shown that, besides β-thujaplicin, there were significant peaks for three diterpenoid compounds, i.e., ferruginol, totarol, hydroxy totarol were detected after matching the spectrum with a library database. The amount of β-thujaplicin was lower compared to other diterpenoid peaks. In *C. cashmeriana*, among the controls β-thujaplicin was detected in NoAgro sample but not in EV samples as it is expected to have some endogenous β-thujaplicin. Compared to the EV sample, OE samples have more β-thujaplicin content, the highest for TpTS1 and TpDOX1. In *C. lusitanica*, the content of β-thujaplicin was high in NoAgro sample and no β-thujaplicin was detected in the overexpressed group (Figure 4.8). Here the metabolite from *C. lusitanica* TpDOX1 overexpressed sample couldn't be extracted due to the lack of sample during preparation. The induction was combined with yeast extract. In *C. lusitanica*, the content of β-thujaplicin was high in NoAgro sample and reduced to overexpressed group (Figure 4.9).

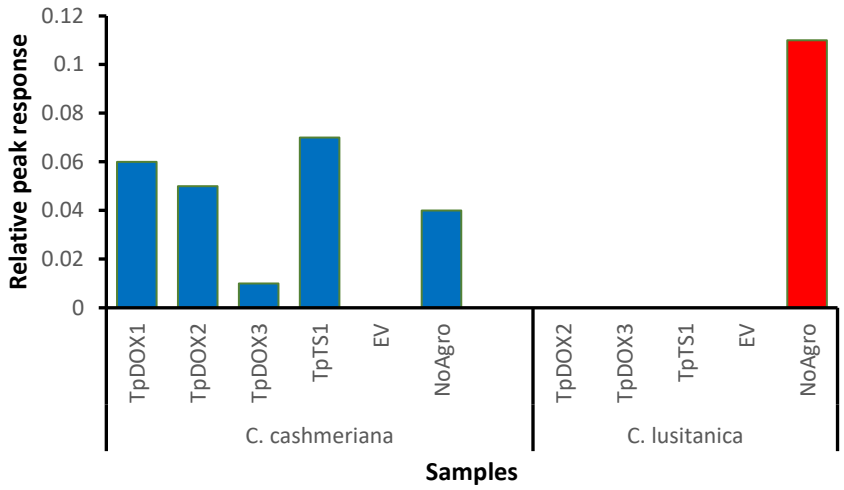


Figure 4.8. Relative peak response for β-thujaplicin content in C.lusitanica and C. cashmeriana cells.

Relative peak response obtained after correcting for internal standard response. Relative peak response = Sample peak area/internal standard peak area.

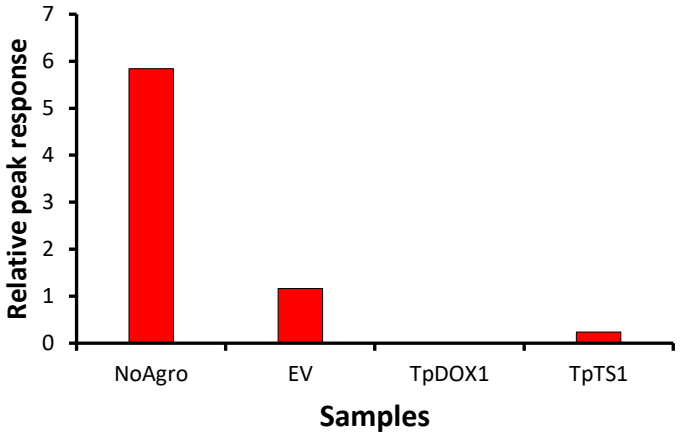


Figure 4.9. Relative peak response for β-thujaplicin content in C.lusitanica cells with yeast extract.

Relative peak response obtained after correcting for internal standard response. Relative peak response = Sample peak area/internal standard peak area.

4.2.8. Stable transformation of dioxygenases

To overcome the debilitating effect of *Agrobacterium*, the experiment was switched from transient to stable transformation. To generate stable transformants, a protocol was

generated for *Agrobacterium*-mediated stable transformation of genes into *C. lusitanica* calli. In addition, to overcome the antibiotic sensitivity, different concentrations of hygromycin were used to test its effect on calli and a standard concentration was selected for the selection of transgenic calli (Figure A.11). Stable transformants couldn't be achieved. As such, during the start of co-cultivation, the calli appeared white (Figure 4.10A) and started to turn red after subsequent transformation into recovery and selection media (Figure 4.10B&C). The calli eventually lost the ability to proliferate in selection media (Figure 4.10D).

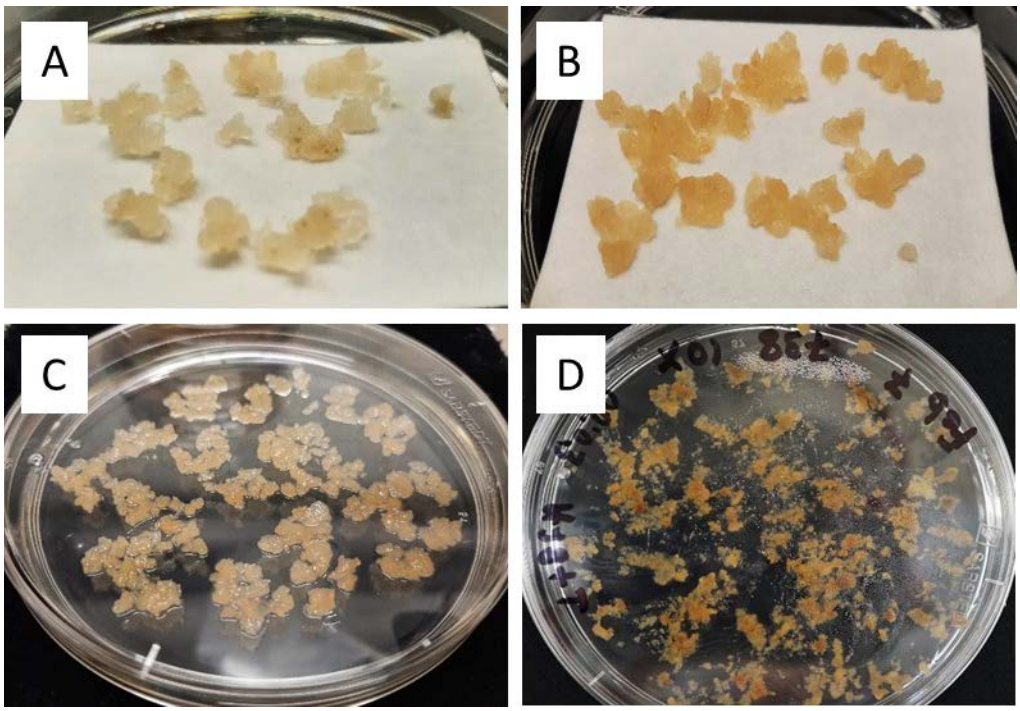


Figure 4.10. Stable transformation in C.lusitanica cells

A) Calli phenotype during co-cultivation media B) Calli phenotype during recovery media C) Calli phenotype during hygromycin selection media D) Calli phenotype during subsequent cell culture.

4.3. Discussion

The genetic basis of tropolone biosynthesis in any plant species is currently unknown. Few papers have been published dealing with the biosynthesis mechanisms of tropolones in plants [231,232,318,319]. Previous studies in *C. lusitanica* cell culture partly revealed some of the biosynthesis steps for the tropolone β-thujaplicin (Figure 1.5,

chapter 1) [223]. The pathway demonstrates that GDP was first converted into terpinolene (steps 1-2), which was further hydroxylated into IME and epoxidized into EMO (steps 2-4) by putative cytochrome P450 (CYP450) monooxygenases [195,223]. The involved genes and enzymes have not been revealed at the molecular level in any plant species. In this study we identified and characterized two genes (*TpTS1*, *TpCS1*) that are upregulated during heartwood formation and whose enzymes produce terpinolene, and therefore are likely to carry out the first dedicated step in β-thujaplicin biosynthesis (See Chapter 2). Studies on one fungal and one bacterial species have shown that a dioxygenase enzyme catalyzes the expansion of the six-carbon ring into the seven-carbon ring characteristic of tropolones. While the dioxygenase enzyme for fungi is dependent on Fe(II) and 2OG as cofactors, the bacterial dioxygenase requires exclusively Fe(II) as a cofactor [231,232]. Based on this evidence, we searched the *T. plicata* heartwood transcriptome for putative dioxygenases of interest.

We identified three putative Fe(II)/2OG dependent dioxygenase genes in a *T. plicata* transcriptome, TpDOX1, 2 and 3, that we think, based on several arguments, are candidates for a role in β-thujaplicin biosynthesis and named them. First, their predicted protein sequences carry the conserved HXD/EXnH motif known to be required for Fe(II) binding and RXS motif for 2OG binding [73,320,321]. The role of Fe (II)/2OG dependent dioxygenases in primary and secondary metabolism has been reported in several species [237,320,322,323]. Secondly, phylogenetic analysis group TpDOX1, 2 and 3 with dioxygenases involved in secondary metabolism and put them in the same subclade as TropC, with a defined role in tropolone-specific ring expansion. Third, they are all expressed in the sap to the heartwood transition zone. Fourth, TpDOX1 is highly upregulated in the transition zone in which the most abundant HW monoterpenoid, β-thujaplicin, is specifically produced.

Differently from terpene synthases, for which well defined substrates are commercially available, we have no such substrates available for TpDOX enzyme assays. In fact, while a radioisotope tracing experiment in *C. lusitanica* cells producing β-thujaplicin identify some intermediates between terpinolene and β-thujaplicin, it failed to identify the immediate precursors to β-thujaplicin (Figure 1.5, chapter 1), possibly because they are short lived. In collaboration with a chemist, a potential immediate precursor for β-thujaplicin named cymorcin was proposed (see full pathway Figure A.12) and we had it commercially synthesized for us. We tested this substrate in *in vitro*

enzyme reactions containing purified TpDOX1, 2 and 3 proteins as well as cofactors Fe(II) and 2OG according to published procedure. Unfortunately, no β-thujaplicin was detected in subsequent GC-MS analyses. To rule out the possibility of defective protein resulting from the purification process we tested an *in vivo* enzyme assay in *E. coli* in which proteins were overexpressed, and enzyme assay was carried out by providing not just cymorcin but also Fe(II) and 2OG. These assays also did not result in β-thujaplicin production. Taken together, the results suggest that cymorcin is either not the right precursor to β-thujaplicin, or that TpDOX1, 2 or 3 do not have a role in β-thujaplicin biosynthesis.

To avoid the need for correct substrate, we turned to *Agrobacterium*-mediated transient over expression of TpDOX1, 2, 3 and TpTS1 in cell cultures of *C.lusitanica* and *C. cashmeriana*, related Cuppressaceae species that differently from *T. plicata*, grow well as calli in our hands. Moreover, *C.lusitanica* cells can produce β-thujaplicin in response to fungal extract [15]. After cocultivation with *Agrobacterium* lines with binary plasmids for TpDOX1, 2, 3 and TpTS1, we observed considerable overexpression, especially with the prime candidate, TpDOX1 (Figure 4.7). Relative to the most relevant negative control, cells exposed *Agrobacterium* and binary plasmid without insert, we observed elevated levels of β-thujaplicin in samples overexpressing TpDOX1, 2 and TpTS1 (Figure 4.8). However, the levels were low, especially relative to the levels of over expressed transcripts, and we did not observe any induction in *C. lusitanica* cells. Since it's possible that suitable endogenous substrate is not produced in these cells, we combined *Agrobacterium* exposure with fungal elicitation in *C.lusitanica*, a treatment shown to induce β-thujaplicin in *C. lusitanica* cell culture [209,229]. We observed elevated β-thujaplicin levels after fungal extract treatment and in the absence of *Agrobacterium*, but this effect was absent in all combinations of fungal extract and Agrobacterium exposure (empty vector, TpDOX1 and TpTS1) (Figure 4.9). Clearly, *Agrobacterium* exposure has a negative impact on β-thujaplicin production, an effect we had seen before (Figure 4.8). In the light of this negative impact, we are unable to draw any conclusion on a potential role of tested enzymes in β-thujaplicin production based on our *in vivo* experiments. To overcome the negative effects of *Agrobacterium*, we have attempted to generate stable transgenic *C. lusitanica* overexpression lines cured from *Agrobacterium* by repeated antibiotic exposure. These attempts have hitherto been

unsuccessful, which is perhaps not surprising, since no transformation procedure has to date been established for this or other related species.

This chapter is exploratory in nature and reveals that there are multiple hurdles to overcome before we can support or reject the hypothesis that any of these TpDOX enzymes are involved in β-thujaplicin biosynthesis. While *in vitro* assays with individual TpDOX proteins and specific substrates would produce the strongest evidence if β-thujaplicin is produced, predicting, and having the correct substrate synthesized is a challenge. We have recently tested another candidate substrate 4(8)-p-menthene-1,2-diol [324] synthesized for us, and again not observed β-thujaplicin production (or any other product) (data not shown). We have considered purifying metabolites from *C. lusitanica* cell lines induced for β-thujaplicin production but lack the expertise and equipment for fractionation and analysis of potential intermediates. From this perspective, *in vivo* assays appear to be the most straightforward approach. Since we have been unable to establish cell cultures from *T. plicata* seedlings, foliage or sapwood, our best choices still are *C. lusitanica* and *C. cashmeriana* cell lines. In light of the observed negative effects of *Agrobacterium* on β-thujaplicin production, the preferred path is to generate stable transgenic lines, possibly using inducible promoters to avoid potential toxicity of involved proteins. The lack of transformants is probably due to a combination of low stable transformation frequency and low survival frequency from hygromycin selection. We have recently started to use a binary plasmid carrying genes that results in transformed cells that produce a red product visible under stereomicroscope. This visual marker may allow us to overcome the above problems.

Chapter 5. An assessment of the effects of stump removal before planting on expression of stress and secondary metabolite biosynthesis genes

5.1. Materials and Methods

5.1.1. Experimental design and sample information

T. plicata samples were collected from 50-year-old trees at an experimental site in Skimikin (50°48′ 32″N and 119°25′16″W), located near Salmon Arm, in the southern interior of British Columbia, Canada. This site has two adjacent 80 m x 160 m blocks, with one stumped by excavator and the other left unstumped, before planting. Each block was divided into 32 plots, each 20 m square (Figure A.13). Single species and two species mixtures were planted in both stump treatments. In this study, cedar trees grown in monoculture and in admixtures with Douglas-fir were randomly sampled from four stumped and six unstumped plots. From stumped plots, monoculture cedar trees were collected from plot no. 4 (5 trees), 12 (5 trees) and 9 (5 trees) whereas cedar trees growing with fir were collected from plot 24 (7 trees). In unstumped plots, monoculture cedar trees were collected from plot 33 (5 trees), 46 (5 trees) and 56 (5 trees) whereas cedar trees growing with fir were collected from plot 37 (5 trees), 42 (5 trees) and 51 (5 trees) (Figure A.13). By using a 12.5mm increment borer, ~8cm wood cores were sampled from fifty-two trees at breast height. These trees were sampled in June (early summer) and same trees were sampled in August (late summer). To recover an adequate quantity of RNA two cores were collected from each individual tree. Tree cores were immediately stored on dry ice and then stored at -20°C laboratory freezer before proceeding with RNA extraction.

5.1.2. RNA extraction

Three different sections i.e., sapwood, sapwood to heartwood transition zone and two rings of heartwood were pooled together for total RNA extraction. These sections were cut into small pieces with chisel and hammer and ground into a fine powder using a freezer mill (SPEX CertiPrep 6750 Freezer Mill, Metuchen, NJ, USA). The extraction procedure followed a modified CTAB protocol as previously described [325]. Briefly,

about 2 g of grounded powder was dissolved in a modified CTAB extraction buffer (2% CTAB; 200 mM Tris, pH 8.0; 50 mM EDTA, pH 8.0; 1 M NaCl; 0.5% activated charcoal; 1.5% polyvinylpolypyrrolidone; and 1.5% β-mercaptoethanol) for 15-30 min at 65°C. The mix was extracted 2 to 3 times with chloroform to sufficiently remove cell debri, protein, polysaccharides, followed by isopropanol-based precipitation of nucleic acids, solubilization of pellet and Li-Cl based precipitation of RNA. The RNA concentration and purity were measured by NanoDrop spectrophotometer (Nanodrop 2000, Thermo Scientific). The integrity of RNA was analyzed by bleach gel electrophoresis [262].

5.1.3. cDNA library preparation, and 3' end mRNA sequencing

To isolate the mRNA from ribosomal RNAs and tRNAs, lexogen's Poly(A) RNA selection kit (Lexogen inc, Austria) was used, which uses Magnosphere oligo(dT) beads that hybridizes to the polyadenylated 3' ends of mRNAs. The concentration of polyA RNA was quantified with Qubit fluorometer by using Qubit RNA High Sensitivity assay kit (Invitrogen). The cDNA libraries were prepared from 1 µg of polyA enriched mRNA using QuantSeq 3' mRNA-Seq Library Prep Kit FWD for Illumina Cat. No. 015.24 (Lexogen, Austria) and multiplexing barcodes for Illumina (i7 index primers 7001-7096; Lexogen, 044.96) with 13 PCR cycles for library amplification. The quantity of the libraries was determined with on a Qubit fluorometer, pooled together for multiplexing, and sent to Genome Quebec at McGill University, Montreal, Canada for sequencing on an Illumina HiSeq 2000 platform.

5.1.4. Quality control and data pre-processing

For June samples, the illumina sequencing generated about 5 million reads of 50 bp lengths for each tree library. The 3'-end reads were trimmed to remove adapter sequences (12 bp), polyA stretch and low-quality tails by the use of BBDuk (https://jgi.doe.gov/data-and-tools/software-tools/bbtools/bb-tools-user-guide/bbduk-guide/). The quality of the reads, before and after trimming, was assessed using FastQC (https://www.bioinformatics.babraham. ac.uk/projects/fastqc/). The trimmed reads were aligned to a primary *T. plicata* wood-derived transcriptome using Bowtie2 and the number of reads per transcript were counted using RSEM [326]. For August samples, the count matrix was obtained from [263]. The count matrices for June and August were used for subsequent statistical analysis.

5.1.5. Identification of differentially expressed genes

The count matrices for June and August were loaded into the limma package in R and a linear model was used to assess the differential gene expression [327]. In this analysis, the differences for sequencing depth and RNA composition were normalized using a TMM normalization method [328]. The raw counts were converted into counts-per-million (CPM) and genes with expression values less than 1 CPM in 6 or more libraries was filtered out. A multidimensional scaling plot based on the log2(CPM) was generated to show relationships between samples. A design matrix was created using R function model.matrix that describes the model to be fitted [329]. The function voom [330] used the above normalisation factor and design matrix to calculate associated weights, ready for linear modelling. The blocking effect for plot was handled using function duplicate correlation and a linear model was fitted using the weights calculated above for each gene. The eBayes function, which performs empirical Bayes shrinkage on the variances, and estimates moderated t-statistics and the associated p-values. Genes with an adjusted P-value ≤ 0.05 were considered significantly differentially expressed and used, respectively. The gene expression patterns were visualized using R function heatmap2 in the gplots package. The functional enrichment analysis of DEGs was performed with GOseq R package [331]. A false discovery rate (FDR) of ≤0.05 was used as a cut-off for the enrichment. The differences of average expression of terpene synthases were tested by linear mixed model analysis. The analysis was done in the JMP 15.0 statistical software. To determine the statistical significance of average expression differs by stump treatment, we used a linear mixed effect model analysis. Here gene expression was included as the response variable, treatment as a fixed effect, and plots as a random effect. The accuracy of the model assumptions was validated by drawing a residual plot. The analysis showed no significance due to an outlier and possible violation of the normality assumption. The analysis was run again without the outliers and showed a significant difference in expression with a p-value of 0.0237. However, we found only 5% of the variation for plots, therefore, in addition, we removed plots as a random effect and performed a nonparametric Wilcoxon test, which is resistant to outliers and doesn’t require any normality assumptions.

5.2. Results

5.2.1. Treatment effects on number of differentially expressed genes (DEGs)

The fungi *Armillaria ostoyae* causes root and stem butt rot in *T. plicata* trees. Studies of a site infected with *A. ostoyae* has shown that removal of stumps after harvest and before planting in 1968 resulted in higher tree survival rate, improved growth, and higher wood density in part because of higher levels of terpenoid extractives [7,332]. Here we aimed at assessing if trees from stumped and unstumped plots differ in gene expression 50 years after the initial treatment, and if so, if gene expression profiles could shed some light on a possible cause of that. In our first analysis, we considered treatment unstumped versus stumped as fixed effects and evaluated the gene expression for monoculture of *T. plicata* (Cw) in stumped versus unstumped plots. This analysis found that in June, a total of 233 DEGs had statistically significant differences (fold change 1) between trees from unstumped and stumped plots, among them 129 and 60 were 1.5 and 2-fold (false discovery rate, FDR <0.05) differentially expressed (Figure 5.1A). In RNA from the same individuals harvested in August, we found a higher number of genes differentially expressed, a total of 1443, 1033 and 623 were statistically significant with fold change 1, 1.5 and 2 (false discovery rate, FDR <0.05) (figure 5.1B). In a second analysis, we considered treatment monoculture *T. plicata* (Cedar west, commonly abbreviated Cw) vs. *T. plicata* mixed with *Pseudotsuga menziesii,* a.k.a. Douglas fir (Df, and a mix referred to as Cwfd), as fixed effects and evaluated the gene expression as above. However, no significant difference in gene expression was found, suggesting mono versus mixed culture had no effect on wood gene expression. Therefore, we increased statistical power for unstumped versus stumped treatment comparison by pooling Cw and Cwfd replicates. We found that the number of DEGs almost doubled in June samples with 449, 224 and 92 genes at 1, 1.5 and 2-fold cut-off levels respectively (Figure 5.1C). Similarly, for August the number increased to 3059, 1791 and 769 for 1, 1.5 and 2-fold (figure 5.1D) (FDR <0.05).

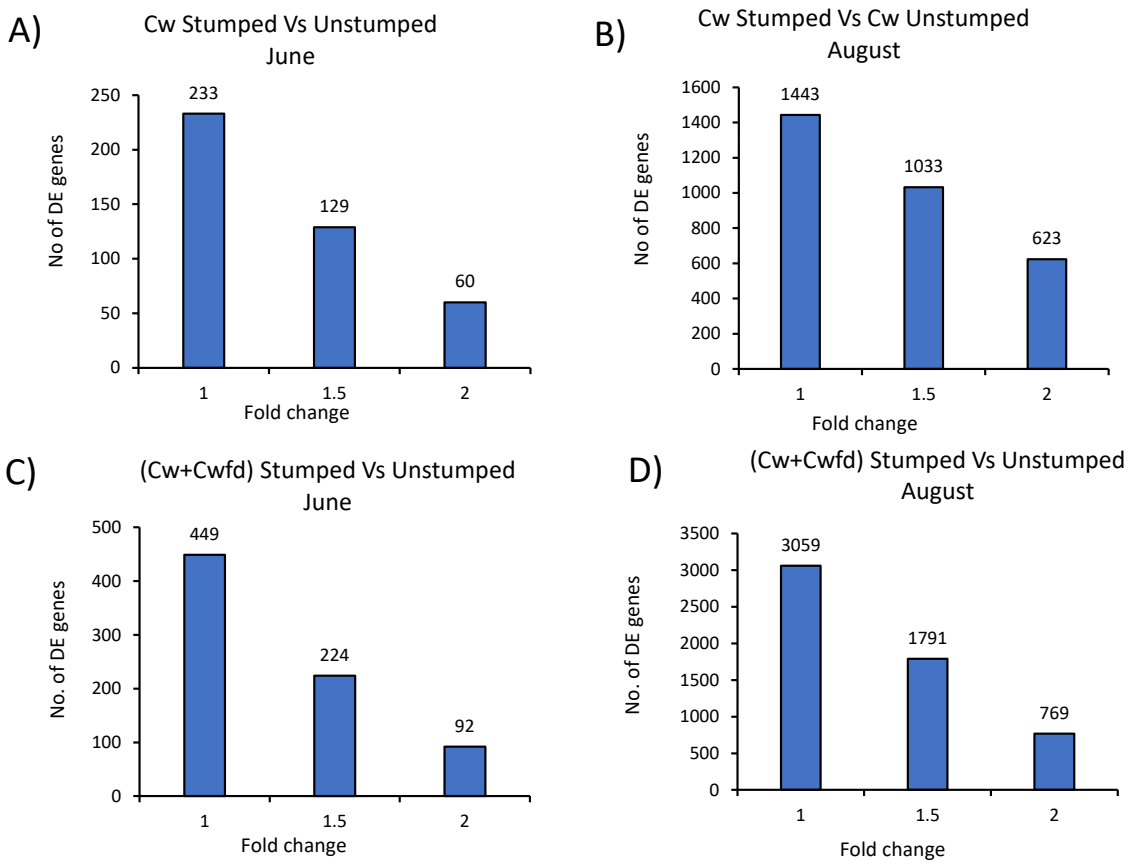


Figure 5.1. Effect of stumped and unstumped pre-treatment on wood gene expression in 50-year-old trees.

Fold-change on x-axis, number of differentially expressed genes on y-axis. Comparison of *T. plicata* (Cw) monoculture in Stumped Vs Unstumped in samples taken in June (A) and August (B). Comparison of *T. plicata* samples pooled from monoculture (Cw) and mixed culture with Douglas fir (Cw+Cwfd) in Stumped Vs Unstumped in samples taken in June (C) and August (D).

5.2.2. Downregulation of stress response in stumped plots vs unstumped plots in August

Gene Ontology (GO) term analysis is a commonly used tool to identify if specific processes are affected by a treatment via comparing the differentially expressed genes to the functional annotation attached to similar genes in a database. The small gene set found to be differentially expressed in wood of trees from stumped versus unstumped plots sampled in June was small and provided just a few GO terms that were enriched. Among the upregulated genes, terms related to DNA-binding transcription factor activity

were overrepresented and in downregulated genes, nitrate assimilation, gibberellin 20-oxidase activity were enriched. The set of differentially expressed in August was larger and provided more GO enriched terms. Among the upregulated terms that were enriched in stumped trees were several related to photosynthesis (Table 5.1). This is surprising considering that the RNA samples come from wood under the bark of mature trees. Either there is residual light reaching the wood, or these genes have plastid functions besides photosynthesis. About 76 genes related to integral component of membrane were enriched and among them, about 10 genes related to cytochrome P450, which are frequently involved in secondary metabolite biosynthesis.

Table 5.1. List of top GO-enriched terms genes 2-fold upregulated in stumped relative to unstumped in August.

GO term category available on GO site, p-value of enrichment of category, number of differentially expressed genes (DEGs) enriched in that category, term for category, and ontology definition. CC=Cellular Component, MF=Molecular Function, BP=Biological Process, FDR=false discovery rate.

Category	P-value	No. DEGs	Term	Ontology	FDR
GO:0009522	2.68E-08	8	photosystem I	CC	0.0002
GO:0009535	7.16E-08	20	chloroplast thylakoid membrane	CC	0.0002
GO:0016168	7.70E-08	8	chlorophyll binding	MF	0.0002
GO:0048544	4.23E-07	10	Recognition of pollen	BP	0.0008
GO:0009765	8.06E-07	5	photosynthesis, light harvesting	BP	0.0013
GO:0018298	1.12E-06	8	protein-chromophore linkage	BP	0.0015
GO:0009523	3.74E-06	7	photosystem II	CC	0.0043
GO:0016021	2.29E-05	76	integral component of membrane	CC	0.0232
GO:0005516	6.82E-05	11	calmodulin binding	MF	0.0615
GO:0010214	8.80E-05	5	Seed coat development	BP	0.0713
GO:0042754	0.0002	3	negative regulation of circadian rhythm	BP	0.1571
GO:0030246	0.0003	13	carbohydrate binding	MF	0.1571

Table 5.2. List of top GO enriched genes 2-fold downregulated in Stumped relative to Unstumped in August.

GO term category available on GO site, p-value of enrichment of category, number of differentially expressed genes (DEGs) enriched in that category, term for category, and ontology definition. CC=Cellular Component, MF=Molecular Function, BP=Biological Process, FDR=false discovery rate.

Category	P-value	No. DEGs	Term	Ontology	FDR
GO:0051082	6.86E-14	21	unfolded protein binding	MF	5.56E-10
GO:0006457	1.82E-13	19	protein folding	BP	7.39E-10
GO:0051259	5.97E-13	12	protein complex oligomerization	BP	1.61E-09
GO:0042542	5.99E-12	14	response to hydrogen peroxide	BP	1.21E-08
GO:0000302	5.59E-11	12	response to reactive oxygen species	BP	9.07E-08
GO:0009116	8.86E-10	9	nucleoside metabolic process	BP	1.20E-06
GO:0043621	1.14E-09	13	protein self-association	MF	1.32E-06
GO:0003824	1.45E-09	9	catalytic activity	MF	1.46E-06
GO:0009408	1.74E-09	16	response to heat	BP	1.57E-06
GO:0045735	1.61E-08	10	nutrient reservoir activity	MF	1.31E-05
GO:0061077	3.81E-07	7	chaperone-mediated protein folding	BP	0.0003
GO:0009644	5.20E-07	9	response to high light intensity	BP	0.0004
GO:0009651	1.89E-06	20	response to salt stress	BP	0.0012
GO:0015250	3.02E-05	5	water channel activity	MF	0.0175
GO:0042807	0.0001	3	central vacuole	CC	0.0555
GO:0046524	0.0002	2	sucrose-phosphate synthase activity	MF	0.1071

To assess the putative functions of the genes upregulated in August in trees growing at unstumped site relative to stumped sites, BLASTX was used to compare their sequence similarity with other known sequences (Table 5.3). Genes related to Abscisic Acid (ABA) biosynthesis such as putative 9-cis-epoxycarotenoid dioxygenases (NCED3), ABSCISIC ACID-INSENSITIVE 5-like protein, abscisic acid receptor were upregulated. Several heat shock proteins (Class I, II & III), transporters such as copper, inositol, iron, magnesium, amino acid, proline, and ammonium, also chaperone protein dnaJ, molecular chaperone regulator, aquaporins all were upregulated. Others stress related genes such as Leucine-rich repeat receptor-like serine/threonine-protein kinase, disease resistance protein, heat stress transcription factor, universal stress protein A-like protein, ankyrin repeat containing protein, ethylene-responsive transcription factor were

upregulated (e-value <0.0, logFC >0.5 and P-value and FDR <0.05). Overall, these multiple processes related to the effects of drought and heat stress show that cells in the wood trees at unstumped sites are more stressed than cells in wood of trees at stumped sites.

Table 5.3. Stress-related genes upregulated in trees from unstumped vs stumped plot in August.
Gene ID in wood transcriptome, description based on top BLASTX match, low E-value indicates high sequence similarity, logFC indicates fold upregulation, P-value of upregulation, and false discovery rate of upregulation (all below 0.05).

GeneID	Description	E-value	logFC	P-value	FDR
Abscisic acid (ABA)					
DN61301c0_g1_i1	E3 Ubiquitin protein ligase	1E-60	0.6554	0.0011	0.0247
DN71010c0_g1_i3	protein phosphatase 2C	0.0	0.6276	1.48E-05	0.0012
DN58498c0_g1_i1	Transcription factor MYB3	1E-19	1.8099	0.0031	0.0460
DN75193c0_g2_i2	9-cis-epoxycarotenoid dioxygenase NCED3	0.0	2.6469	8.90E-07	0.0002
DN55257c0_g1_i4	9-cis-epoxycarotenoid dioxygenase NCED3	0.0	1.8720	0.0004	0.0119
DN52390c0_g1_i1	ABSCISIC ACID-INSENSITIVE 5-like protein	3E-35	0.8304	2.78E-09	2.61E-06
DN44293c0_g1_i3	Abscisic acid receptor	2E-90	0.6916	0.0019	0.0341
DN40374c1_g1_i1	Dehydration-responsive element-binding protein	6E-31	2.2039	3.87E-05	0.0023
DN47599c1_g1_i4	ABSCISIC ACID-INSENSITIVE 5-like protein	7E-41	0.9921	0.0027	0.0431
DN80920c0_g8_i2	Putative dehydration-responsive element-binding protein	5E-20	1.2683	0.0026	0.0417
DN47421c0_g2_i4	Mitogen-activated protein kinase	2e-60	2.1561	0.0017	0.0313
Reactive oxygen species (ROX)					
DN76808c1_g4_i1	class I heat shock protein	1E-23	0.7809	0.0006	0.0169
DN74323c0_g3_i6	class I heat shock protein	6E-39	1.4679	0.0001	0.0054
DN74723c0_g2_i2	class II heat shock protein	2E-49	2.1047	9.35E-07	0.0002
DN75575c0_g1_i8	class II heat shock protein	2E-38	1.7246	3.04E-07	8.23E-05
DN50855c1_g2_i6	L-ascorbate peroxidase	0.0	0.4754	3.31E-06	0.0004
DN44804c2_g1_i4	Chromophore lyase	1E-12	0.3958	1.26E-05	0.0011
DN81519c1_g2_i6	Epoxide hydrolase B	5E-37	0.3406	0.0006	0.0161
DN51477c0_g1_i1	Serine/threonine-protein kinase	0.0	1.1270	2.53E-05	0.0017
DN55993c2_g1_i6	Heat shock protein	0.0	1.5223	0.0005	0.0148
DN54718c0_g2_i10	Small heat shock protein	1e-56	2.0956	0.0009	0.0227
DN55993c1_g1_i13	Heat shock protein	0.0	1.2838	8.31E-06	0.0008
DN75308c1_g2_i1	class I heat shock protein	6e-67	1.0688	1.09E-05	0.0009
DN53946c1_g1_i6	Heat shock cognate 70 kDa protein	0.0	1.3591	0.0008	0.0199
DN38261c0_g1_i2	class III heat shock protein	5e-22	1.3122	0.0005	0.0141
DN39593c0_g1_i3	Small heat shock protein	2e-61	1.6399	8.77E-07	0.0002
DN66010c0_g1_i1	class IV heat shock protein	6e-51	1.2562	2.74E-05	0.0018
DN53300c0_g1_i3	class II heat shock protein	2e-39	1.9849	9.35E-05	0.0044
DN39726c0_g1_i1	Small heat shock protein	8e-34	1.1284	8.05E-05	0.0039
DN55993c1_g1_i6	Heat shock protein	0.0	2.3870	1.16E-05	0.0010
DN50095c1_g1_i6	class I heat shock protein	4e-43	1.9489	0.0003	0.0093

DN75018c0_g2_i2	heat shock protein	4e-23	1.5168	2.43E-10	5.37E-07
DN75308c0_g1_i1	class I heat shock protein	2e-63	1.7244	0.0032	0.0475
DN53300c0_g1_i1	class II heat shock protein	2e-49	1.6749	1.45E-06	0.0002
DN55993c2_g1_i9	Heat shock protein	1e-14	1.4446	6.61E-06	0.0007
DN75575c0_g1_i9	class II heat shock protein	2e-38	1.8482	1.61E-05	0.0013
DN50095c1_g1_i8	class I heat shock protein	1e-39	1.4393	0.0029	0.0454
DN80127c1_g1_i4	class I heat shock protein	2e-62	1.7456	2.79E-06	0.0004
DN76769c1_g1_i10	Small heat shock protein	1e-47	1.7218	1.54E-05	0.0012
DN54718c0_g2_i9	Small heat shock protein	8e-56	1.7900	1.01E-06	0.0002
DN53560c0_g1_i15	class IV heat shock protein	2e-47	1.2364	3.35E-05	0.0021
DN53300c0_g1_i13	class II heat shock protein	1e-33	1.6962	7.39E-07	0.0002

Transporters

DN47619c0_g1_i3	Probable plastidic glucose transporter	1E-17	0.9407	5.61E-05	0.0030
DN49182c0_g1_i1	Probable polyamine transporter	1E-16	1.8312	0.0010	0.0231
DN45312c0_g2_i2	Copper transporter	2E-12	0.7851	0.0018	0.0331
DN57713c0_g1_i18	Probable inositol transporter	2E-16	0.8788	0.0031	0.0466
DN56626c0_g2_i3	Bidirectional sugar transporter SWEET1	8E-48	1.0409	0.0017	0.0324
DN70284c0_g1_i6	Magnesium transporter	2e-14	1.0445	1.65E-10	4.54E-07
DN76495c0_g3_i2	Vacuolar iron transporter	2e-08	0.8522	0.0003	0.0103
DN69687c0_g1_i1	Amino acid transporter	3e-10	0.4792	0.0028	0.0434
DN19225c0_g1_i1	Magnesium transporter	3e-10	0.6963	0.0012	0.0249
DN66271c1_g2_i1	Inorganic phosphate transporter	0.0	0.9374	0.0014	0.0286
DN76941c2_g2_i11	Probable proline transporter	0.0	0.8574	0.0027	0.0429
DN75967c0_g2_i16	Ammonium transporter	0.0	1.0068	0.0030	0.0458
DN81234c1_g2_i12	Bidirectional sugar transporter SWEET1	1E-96	2.1172	3.29E-07	8.37E-05
DN79503c6_g2_i4	Probable polyol transporter	2E-14	1.8852	5.86E-06	0.0006
DN57351c0_g1_i1	Choline transporter protein	0.0	0.4876	1.54E-05	0.0012
DN59064c1_g1_i2	Oligopeptide transporter	2e-49	1.0902	0.0030	0.0459
DN55604c1_g1_i2	ABC transporter C family member	0.0	1.1611	0.0006	0.0165
DN81538c0_g1_i1	Oligopeptide transporter	0.0	1.3783	0.0020	0.0357
DN63991c0_g1_i1	metal-nicotianamine transporter	0.0	1.8046	0.0002	0.0077
DN77206c0_g2_i1	Molybdate-anion transporter	1e-62	1.4734	0.0012	0.0261

Chaperones

DN77486c7_g1_i9	Heat shock protein	1e-11	1.3425	2.31E-05	0.0016
DN79431c1_g1_i6	Chaperone protein dnaJ	9e-19	1.3954	0.0011	0.0242
DN70199c0_g1_i1	Chaperone protein DnaJ	9e-17	1.6632	1.49E-05	0.0012
DN79566c2_g5_i1	Chaperone protein ClpB3	0.0	0.3068	0.0006	0.0165
DN74460c1_g6_i1	Chaperone protein dnaJ 8	9e-22	0.688	0.0002	0.0067
DN76410c0_g1_i15	20 kDa chaperonin	7e-10	0.688	1.78E-08	1.04E-05
DN78266c1_g1_i6	Chaperone protein dnaJ	8e-15	1.065	0.0001	0.0055
DN78266c1_g1_i16	Chaperone protein dnaJ	3e-16	1.1305	0.0004	0.0116
DN63771c0_g1_i4	20 kDa chaperonin	3e-94	1.0510	1.62E-06	0.0003
DN74936c0_g1_i1	Chaperone protein dnaJ	2e-14	1.5574	0.0023	0.0386
DN72683c0_g1_i4	BAG family molecular chaperone regulator	2e-12	1.5769	0.0002	0.0077
DN48197c0_g1_i1	DnaJ homolog subfamily B member	1e-15	1.0707	1.24E-13	3.31E-09
DN54764c0_g1_i2	Chaperone protein dnaJ	2e-22	1.3279	7.36E-05	0.0037
DN45170c0_g1_i1	BAG family molecular chaperone regulator	7e-14	1.5854	0.0007	0.0178
DN50458c0_g1_i5	DnaJ homolog	1e-37	1.6361	0.0017	0.0321

DN45645c3_g2_i8	Chaperone protein DnaJ	9e-16	1.6299	1.32E-07	4.47E-05
DN48002c2_g1_i3	Chaperone protein dnaJ	3e-23	1.1568	6.42E-06	0.0007
DN49925c0_g1_i4	Chaperone protein dnaJ	2e-20	1.1177	0.0005	0.0134

Aquaporin

DN52929c0_g2_i1	Aquaporin PIP2-7	2e-73	1.1247	2.13E-05	0.0015
DN76906c6_g1_i4	Aquaporin TIP1-1	6e-11	1.4818	1.27E-05	0.0011
DN76715c1_g1_i5	Aquaporin TIP2-1	1e-10	1.2025	3.87E-06	0.0005
DN55057c0_g1_i1	Probable aquaporin PIP2-8	3e-14	1.7030	0.0002	0.0081
DN76715c1_g3_i2	Aquaporin TIP2-1	9e-10	1.5326	0.0006	0.0156

Others

DN38173c0_g1_i2	Zinc finger AN1 domain-containing stress-associated protein	2e-38	1.6737	0.0021	0.0369
DN56061c1_g1_i5	TMV resistance protein	2e-14	1.3778	0.0015	0.0298
DN50517c1_g2_i6	Galactinol synthase	4e-53	1.7365	8.61E-05	0.0042
DN57508c0_g2_i2	Leucine-rich repeat receptor-like serine/threonine-protein kinase	0.0	1.1754	0.0011	0.0241
DN74740c0_g1_i1	Universal stress protein A-like protein	9e-06	1.4982	0.0003	0.0112
DN66966c0_g1_i2	NDR1/HIN1-like protein	6e-27	1.0430	0.0028	0.0439
DN53856c0_g1_i1	Universal stress protein Sll1388	1e-08	1.0659	0.0001	0.0060
DN52128c1_g1_i5	Disease resistance protein	1e-12	1.5561	2.54E-05	0.0017
DN74588c0_g1_i15	Galactinol synthase	2e-58	1.7928	3.03E-05	0.0019
DN74310c1_g3_i3	Disease resistance protein	8e-18	1.5971	0.0003	0.0110
DN51736c1_g3_i9	Heat stress transcription factor	3e-85	1.6283	0.0011	0.0244
DN68312c0_g1_i1	NDR1/HIN1-like protein	2e-07	2.0529	0.0002	0.0079
DN78518c1_g4_i1	Defensin	1e-25	2.3057	5.28E-06	0.0006
DN76698c0_g2_i5	Heat stress transcription factor	3e-51	1.5387	2.01E-05	0.0015
DN53070c2_g1_i1	Wound-induced protein	1e-79	3.8801	0.0002	0.0087
DN49773c0_g1_i14	Universal stress protein A-like protein	2e-06	2.2501	0.0009	0.0222
DN49479c1_g2_i2	Probable disease resistance protein	5e-16	1.2091	0.0030	0.0458
DN53099c0_g3_i3	Ethylene-responsive transcription factor	9e-46	3.1919	0.0003	0.0090
DN51824c1_g1_i4	Ankyrin repeat-containing protein	1e-19	1.0416	0.0009	0.0226
DN56883c0_g1_i6	LRR receptor-like serine/threonine-protein kinase	1e-50	1.0467	0.0016	0.0312
DN74504c2_g12_i1	Ethylene-responsive transcription factor	1e-33	2.6517	1.48E-05	0.0012
DN52894c0_g2_i3	Ethylene-responsive transcription factor	4e-27	1.5791	0.0007	0.0179
DN50341c1_g2_i4	Disease resistance protein	5e-06	1.2353	0.0003	0.0090
DN58645c1_g1_i2	Disease resistance protein	1e-18	2.1493	5.93E-07	0.0001
DN76027c0_g1_i2	Ethylene-responsive transcription factor	7e-22	1.2011	0.0016	0.0308
DN80376c1_g1_i10	Disease resistance protein	2e-18	1.3855	8.67E-05	0.0042
DN78947c0_g1_i2	Protein STRICTOSIDINE SYNTHASE-LIKE	2e-84	2.0877	0.0025	0.0405
DN77522c3_g2_i1	Universal stress protein	5e-14	1.0684	0.0012	0.0255
DN55219c1_g1_i1	Disease resistance protein	1e-07	2.5413	0.0001	0.0054
DN58645c1_g1_i3	Disease resistance protein	7e-29	1.3313	0.0019	0.0343
DN79901c0_g1_i2	Ethylene-responsive transcription factor	2e-35	2.0575	0.0018	0.0328
DN74823c0_g7_i1	Hydrophobic protein LTI6A	1e-22	1.5212	0.0004	0.0121
DN54768c0_g1_i1	Universal stress protein	4e-65	1.0942	0.0003	0.0107
DN54774c0_g1_i6	Heat stress transcription factor	5e-48	1.3242	0.0004	0.0121

5.2.3. Upregulation of terpene synthase-encoding transcripts in wood of stumped trees

Trees growing at stumped plots have shown over decades to have higher wood density and higher levels of heartwood terpenoid extractives. To assess if this effect is still ongoing, we identified transcripts with the terpene synthase annotation in the combined libraries. We found that their combined expression was higher in wood of stumped trees than in unstumped trees in the dataset sampled in August (Figure 5.2).

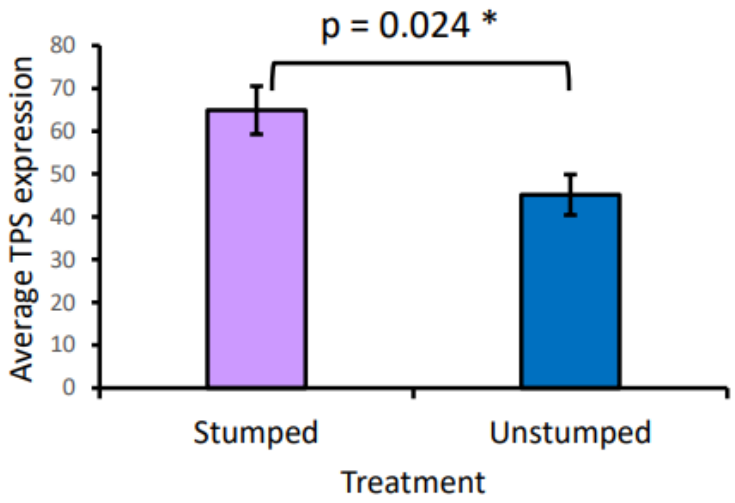


Figure 5.2. Average gene expression of terpene synthases (TPSs) in stumped and unstumped plots.

A linear mixed effect model was used to test the differences in gene expression. Average expression of TPSs shows significant difference between stumped and unstumped with p-value 0.024.

Since we have identified mono and sesquiterpene synthases that produce defined terpenes (chapter 2), we also assessed if their specific transcripts were affected. Figure 5.3 shows that transcripts of the identified terpinolene, 3-carene and elemol synthases were expressed at significantly higher levels in wood of trees from stumped plots relative to trees from unstumped plots.

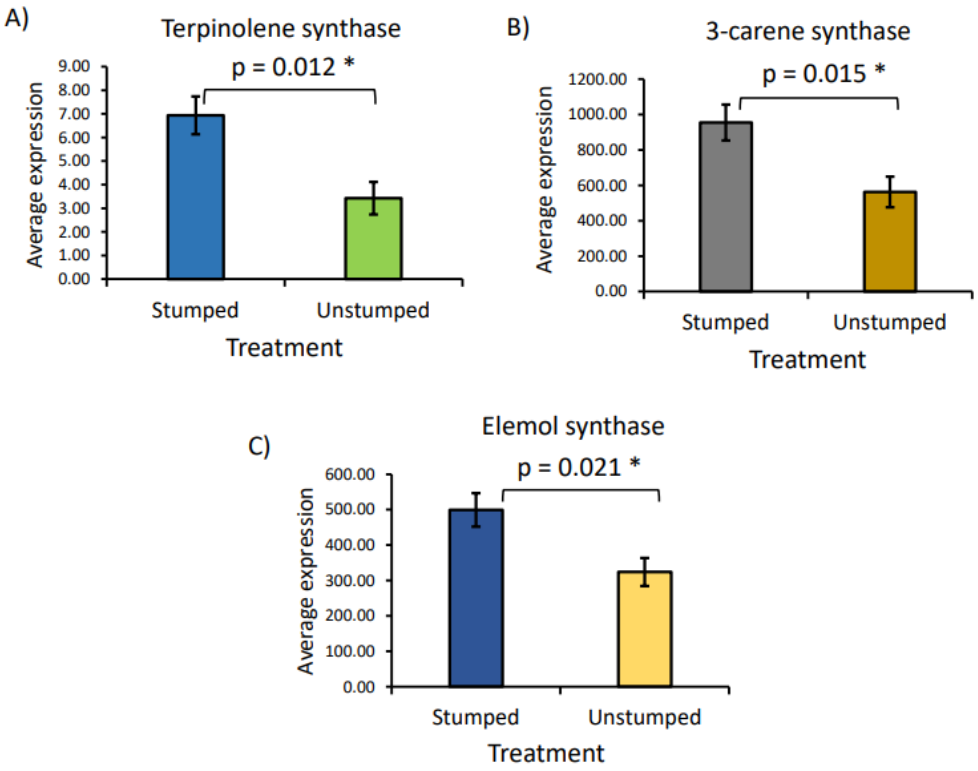


Figure 5.3. Average gene expression of terpinolene, 3-carene and elemol synthases in stumped and unstumped plots.

A linear mixed effect model was used to test the differences in gene expression. The difference of the average expression of terpinolene (A), 3-carene (B) and elemol (C) synthases in stumped and unstumped plots are significant with p-value <0.05.

5.2.4. Climate data from June and August sampling times

We sampled the same trees in June and August because active growth is likely to occur in June, and heartwood extractives are likely to be produced in August. Moreover, the sampled site frequently experiences drought stress in August. Figure 5.4 shows climate data collected at a weather station about 20 km away from the sampled site in the same year. It shows that average temperature was several degrees higher in August than in June, and that precipitation in August was 37% of that in June. Moreover, we did not observe any obvert signs of drought stress in June, whereas treetops showed extensive flagging due to lack of turgor pressure in August.

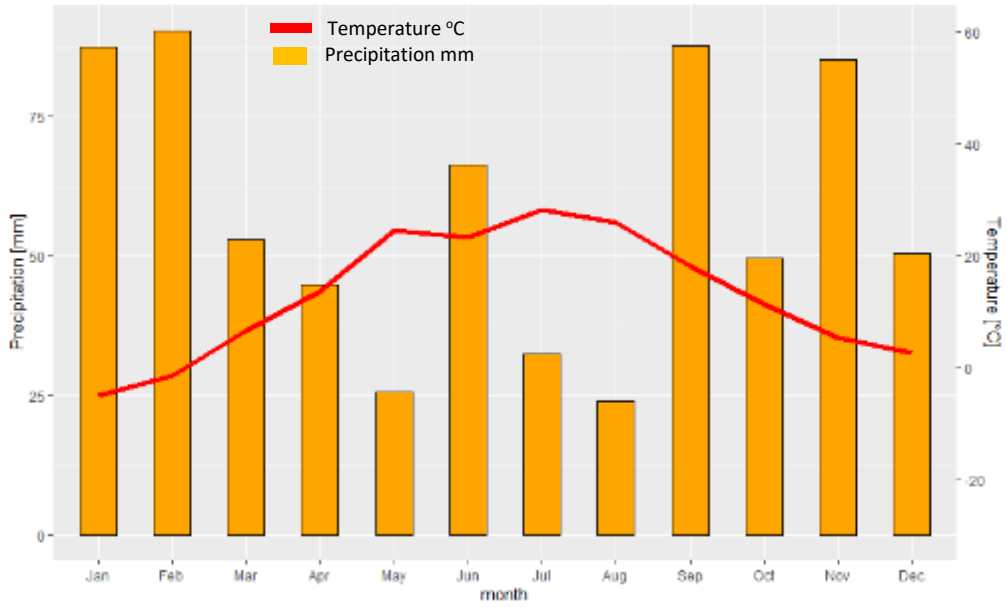


Figure 5.4. Average temperature and precipitation at the Silver Creek weather station in Silver Creek, British Columbia for the year 2018.

The weather station, maintained by CCN, is ~20 km from the sampling site, and data were obtained from a publicly available dataset at [333]. Figure recreated from data obtained from [263].

5.2.5. Differential gene expression in wood of the same trees sampled in June and August

We compared the expression of wood taken from the same trees in June and August to assess the extent of differential gene expression between them. Visualization of DEGs at >2 fold showed that a total of 5661 genes were differentially expressed at significant levels (FDR <0.05), with 2714 upregulated and 2947 downregulated in August relative to June (Figure 5.5).

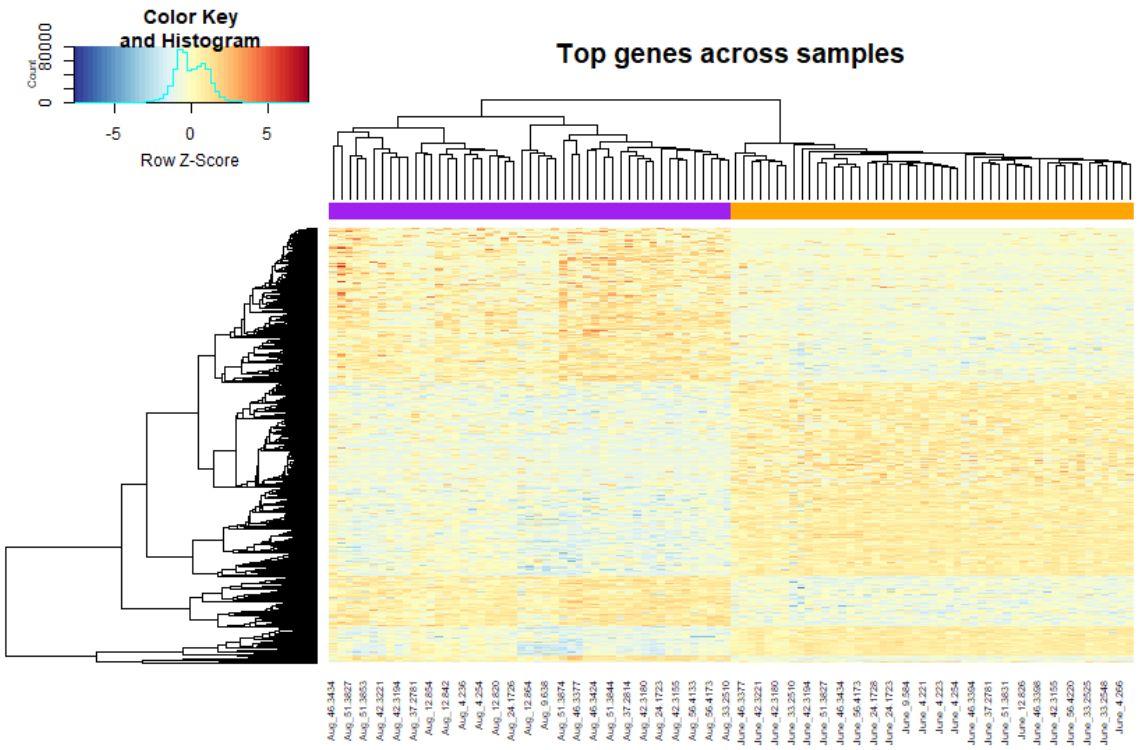


Figure 5.5. Heatmap showing the relative expression of the differentially expressed genes in pairwise comparisons between June and August.

The columns represent DEGs from individual trees, with purple bar indicating August sampling treatment group, and orange bar June sampling treatment group, with a p-value cut-off at 0.05 and a fold change cut-off at 2-fold. Rows indicate individual DEGs, with branches and sub-branches representing genes with similar profiles of differential expression. Yellow to red indicate upregulation, purple to blue indicate down regulation. Top genes across samples = 2-fold differentially expressed genes.

Gene ontology (GO) functional enrichment analysis of the upregulated genes indicated that GO terms such as DNA binding transcription factor activity, defense response, terpene synthase activity, response to wounding, nutrient reservoir activity was significantly overrepresented in August (Table 5.4). Whereas in June, GO terms related to cell wall biogenesis, cell wall organization, apoplast, microtubule binding, microtubule motor activity was significantly overrepresented with an FDR <0.05 (Table 5.5). The enrichment for cell wall biosynthesis GO terms in June indicate active growth whereas the many drought and other abiotic stress-related terms enrichment in August indicate that drought stress was present also in living wood cells. Moreover, the enrichment for the terpene biosynthesis term, with 20 genes expressed at higher levels in August than in June, is constituent with heartwood formation in August.

Table 5.4. Top 10 GO enriched terms with transcripts 2-fold upregulated in August relative to June.
GO term categories available on the GO site, p-value of enrichment of category, number of differentially expressed genes (DEGs) enriched in that category, term for category, and ontology definition. CC=Cellular Component, MF=Molecular Function, BP=Biological Process, FDR=false discovery rate.

Category	Overrepresented p-value	No. of DE in category	term	ontology	Over represented FDR
GO:0003700	1.62E-12	160	DNA binding transcription factor activity	MF	1.32E-08
GO:0006952	9.48E-11	133	Defense response	BP	3.87E-07
GO:0010333	4.07E-10	20	Terpene synthase activity	MF	1.11E-06
GO:0008194	1.35E-09	32	UDP glycosyltransferase activity	MF	2.76E-06
GO:0009611	6.20E-09	49	Response to wounding	BP	8.59E-06
GO:0005576	6.31E-09	118	Extracellular region	CC	8.59E-06
GO:0045735	2.37E-08	22	Nutrient reservoir activity	MF	2.77E-05
GO:0016208	9.92E-08	17	AMP binding	MF	0.0001
GO:0042742	1.71E-07	106	Defense response to bacterium	BP	0.0002
GO:0005886	2.54E-07	349	Plasma membrane	CC	0.0002

Table 5.5. Top 10 GO enriched genes 2-fold upregulated in June relative to August.
GO term categories available on the GO site, p-value of enrichment of category, number of differentially expressed genes (DEGs) enriched in that category, term for category, and ontology definition. CC=Cellular Component, MF=Molecular Function, BP=Biological Process, FDR=false discovery rate.

Category	Over represented p-value	No. of DE in category	term	ontology	Over represented FDR
GO:0009834	2.97E-23	31	Plant type secondary cell wall biogenesis	BP	2.42E-19
GO:0005874	5.02E-14	53	microtubule	CC	2.05E-10
GO:0048046	1.64E-12	102	apoplast	CC	4.48E-09
GO:0005576	7.71E-12	135	Extracellular region	CC	1.57E-08
GO:0071555	1.06E-11	69	Cell wall organization	BP	1.72E-08
GO:0008017	5.14E-11	41	Microtubule binding	MF	7.00E-08
GO:0003777	6.90E-10	21	Microtubule motor activity	MF	7.05E-07
GO:0005871	6.90E-10	21	Kinesin complex	CC	7.05E-07
GO:0007018	1.22E-09	21	Microtubule based movement	BP	1.11E-06
GO:0045492	1.90E-09	11	Xylan biosynthetic process	BP	1.55E-06
GO:0005618	3.15E-08	88	Cell wall	CC	1.98E-05
GO:0030244	7.23E-06	17	Cellulose biosynthetic process	BP	0.0030

5.3. Discussion

To date the effect of stump removal before planting at the described site has been shown to increase survival, growth, and wood density and reduce rot decay, [7,244,245,247,248,332], with the original intention of removing substrate for growth of *A. ostoyae* that had caused extensive root and butt rot at this site before. The cause of the observed beneficial effects is far from fully understood and we test if gene expression analyses of wood, including cambium forming xylem, differentiating xylem, living xylary ray and axial parenchyma cells, and finally, in the transition zone in which the ray cells produce the secondary metabolites before they die. The secondary metabolites, primarily lignans and terpenoids, in turn contribute to both rot resistance and wood density. Since wood growth occur primarily in spring and early summer, whereas secondary metabolites are produced late summer and fall and suspected drought typically occurs in August, we sampled the same 52 trees in June and August. While RNA-seq studies are rarely used on field-grown mature trees because of extensive biological variation, and a risk for identification of DEGs by chance among the many thousands of assessed genes. We found, however, a surprisingly large number of genes that were differentially expressed in wood from trees grown at stumped and unstumped sites when sampled in August (but not in June) and when comparing expression in all sampled trees in June and August. Moreover, large portions of identified gene sets matched the physiological expectations indicating that relevant genes were identified. In particular, genes upregulated in June compared to August were enriched for GO terms and a large portion of genes related to cell wall biosynthesis, consistent with expected cell growth in June. Vice versa, genes upregulated in August were enriched for many GO terms related to stress, in particular drought stress, which is consistent with the drought we observed at time of sampling as well as climate data collected nearby. In the comparison of wood gene expression between trees planted at sites where stumps had or had not been removed before planting 50 years ago, the most noticeable GO term enrichment was for terms related to abiotic stress in wood from unstumped sites. With respect to wood terpene content, related to both rot resistance and wood density, we found that 20 genes with annotations related to terpene biosynthesis as well as a set of cytochrome P450 genes were significantly upregulated in August versus June, and that wood from trees at stumped sites had higher expression of terpene synthases, including candidates for the first committed step in biosynthesis of

the most abundant terpenoid, β-thujaplicin (chapter 2). Since a treatment that took place 50 years ago is unlikely to have a direct effect on gene expression - terpene synthases or drought response genes - we propose a model in which a sequence of events may explain observed differences (Figure 5.6). Previous data show that most trees growing in unstumped plots had root disease, suggesting root damage could limit water and mineral uptake and thereby photosynthesis and growth [12]. In addition, the trees with reduced water uptake would experience more drought stress during late-summer drought spells, frequently occurring in the region, in line with the observed higher levels of drought-related GO terms found to be upregulated in wood from trees on unstumped sites relative to stumped sites and also in line with the extensive enrichment of drought and abiotic stress terms in samples collected in August compared to June. Higher growth and wood density in trees at stumped can be attributed to better access to water and minerals in general. A key component of heartwood density is the amount of extractives it contains, and unpublished data show higher levels of terpenoid extractives in wood of trees grown on stumped sites (Cruikshank, in manuscript). Our results show that (1) terpene biosynthesis is highly active in August relative to June, and (2) that terpene synthases are expressed at higher levels in wood of trees on stumped sites in August, (3) which coincides with drought stress. Thus, it stands to reason that higher wood density and extractive levels are due to less negative impacts of drought stress at this time of the year. In addition to direct effect on gene expression, drought stress likely has an impact on accumulation of starch in amyloplasts stored in ray and axial parenchyma and used up to produce and secrete the large amounts of secondary metabolites in the transition zone (data not shown).

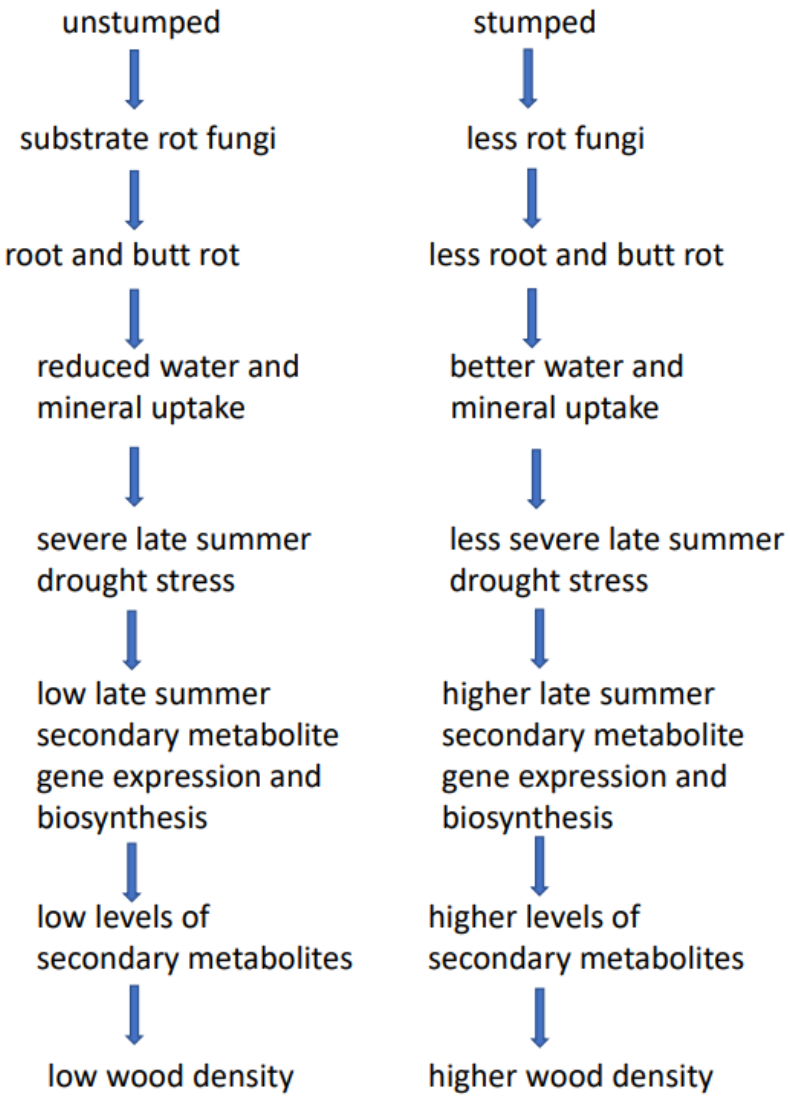


Figure 5.6. Model of sequence of events leading to higher wood density in trees growing on stumped sites.

It proposes that frequently occurring late summer droughts are less severe in trees at stumped sites due to better water uptake, which has a direct impact on terpene biosynthesis occurring at the same time. This model explains how a silvicultural treatment 50 years ago can have effects on gene expression in current time.

Summary Statement

The genetic basis of secondary metabolite biosynthesis in developing heartwood is not understood or understood well in tree species. This may appear surprising considering the high value of heartwood in many species. On the other hand, forest tree species cannot readily be genetically dissected by mutant screens and phenotype analyses as has been done for a selected few annual species with short life cycles. In addition, the process of heartwood formation occurs deep into the wood and is done by a selected few living parenchyma cells among dead tracheids and fibers. Here we used a protocol that allowed recovery of RNA from the region of heartwood in combination with large-scale RNA sequencing to identify genes upregulated during heartwood formation in *T. plicata*. Among them, we focused on transcripts with high sequence similarity to known terpene synthases because these enzymes carry out the first dedicated steps in the production of specific terpenoids. We identified two genes, and corresponding enzymes TpTS1 and TpCS1, that produce terpinolene, the first predicted step in the biosynthesis of the tropolone β-thujaplicin. Moreover TpCS1 produces 3-carene as the most abundant product. None of these monoterpenes have been found at high levels in *T. plicata* heartwood, and they are likely to be converted into β-thujaplicin and other tropolones that are abundant in *T. plicata* heartwood. The identification of a sesquiterpene synthase that produces elemol and diterpene synthases that produce normal-copalyl diphosphate, an intermediate for diterpene biosynthesis, and the diterpenes sandaracopimaradiene or *syn*-stemod-13(17)-ene depending on the substrate, and levopimaradiene. These diterpenes are and their derived acids and alcohols are known to occur in wood resin of other conifer species, in which they are implicated in both rot and insect resistance based on their toxicity. The sesquiterpene elemol is likewise implicated in termite resistance in another species. None of these compounds or their derivatives have to our knowledge been identified in *T. plicata* heartwood. Since existing literature focus on β-thujaplicin (and related tropolones) for discussion and selection efforts towards improved heartwood rot resistance, but the tropolones are largely or entirely gone from older heartwood, these compounds should be considered for durable resistance. Diterpene derivates are known to remain in heartwood for hundreds of years, and although initially produced at lower levels than tropolones, their role in rot resistance could be more relevant than the fleeting tropolones. Taken together, these identified genes, enzymes and their products

constitute the first and fairly large step forward in the understanding of terpene biosynthesis during heartwood formation in *T. plicata*, and also a substantial contribution to the field as a whole. We were unable to show that identified DOX genes carry out the ring expansion characteristic of tropolones, and revised approaches have to be taken to prove or disprove such functions. Our transcriptome also contains candidates for various modifications of identified terpenes into final compounds stored in the heartwood. Strictly speaking, the functions of genes and proteins based on enzymatic activity do not necessarily need to be shown to be useful for marker assisted selection. This study demonstrates that relevant genes and enzyme activities have been found. It should be possible to proceed to identification of associations between variants in these genes and content of corresponding terpene metabolites, and when found, be used as genotype markers for phenotype prediction in young trees, circumventing the need for waiting for 15-20 years before levels of individual terpenoids can be quantified. Moreover, heartwood terpene content may be a much more limited quantitative trait than for example growth and drought resistance, and therefore may not require unaffordable genomic selection to be integrated into a breeding program with limited resources. Finally, we were provided with the opportunity to test if RNA-seq could be used to shed some light on how a silvicultural treatment, removal of stumps from a root and butt rot-diseased site, that took place some 50 years ago could still have an effect on growth and wood quality. Based on gene expression results we could develop a model that is both plausible and provide testable hypotheses to further the understanding of the observed differences, and also help define further applications of this beneficial treatment.

References

1. Chedgy, R. Secondary Metabolites of Western Red Cedar (Thuja Plicata): Their Biotechnological Applications and Role in Conferring Natural Durability. *L. Lambert Acad. Publ. Saarbrücken, Ger.* **2010**, 7–8.

2. Sturrock RN, Braybrooks AV, R.P. Decay of Living Western Redcedar: A Literature Review. *Victoria Nat. Resour. Canada, CFS, Can. Wood Fibre Centre. Inf.* **2017**, *Report FI*.

3. Minore, D. Western Redcedar—A Literature Review. *Gen. Tech. Rep.* **1983**, *PNW-150*, doi:10.2307/3533673.

4. Brodie, L.C.; Harrington, C.A. Response of Western Redcedar to Release and Fertilization in a Mixed-Species Stand. *USDA For. Serv. - Gen. Tech. Rep. PNW-GTR* **2010**, 139–144.

5. Antos, J.A.; Filipescu, C.N.; Negrave, R.W. Ecology of Western Redcedar (Thuja Plicata): Implications for Management of a High-Value Multiple-Use Resource. *For. Ecol. Manage.* **2016**, *375*, 211–222, doi:10.1016/j.foreco.2016.05.043.

6. Klinka, K.; Brisco, D. *Silvics and Silviculture of Coastal Western Redcedar: A Literature Review*; 2009; Vol. 11; ISBN 9780772661104.

7. Cruickshank, M.G.; Filipescu, C.N.; Sturrock, R.N. The Effect of Stump Removal and Tree Admixture on Butt Decay Incidence, Damage and Wood Density in Western Redcedar. *Can. J. Plant Pathol.* **2018**, *40*, 368–377, doi:10.1080/07060661.2018.1496143.

8. Gregory, C.; McBeath, A.; Filipescu, C. *An Economic Assessment of the Western Redcedar Industry in British Columbia*; Technical Report, Information report FI-X-017; Natural Resources Canada, Canadian Forest Service, Canadian Wood Fibre Centre: Victoria, BC, Canada, p. 38, 2018; ISBN 9780660240947.

9. Hebda, Richard J.; Mathewes, R.W. Holocene History of Cedar and Native Indian Cultures of the North American Pacific Coast Author (s): Richard J . Hebda and Rolf W . Mathewes Published by : American Association for the Advancement of Science Stable URL : Https://Www.Jstor.Org/Stable/169. **1984**, *225*, 711–713.

10. Morris, P.I.; Stirling, R. Western Red Cedar Extractives Associated with Durability in Ground Contact. *Wood Sci. Technol.* **2012**, *46*, 991–1002, doi:10.1007/s00226-011-0459-2.

11. Celedon, J.M.; Bohlmann, J. An Extended Model of Heartwood Secondary Metabolism Informed by Functional Genomics. *Tree Physiol.* **2018**, *38*, 311–319, doi:10.1093/treephys/tpx070.

12. Buchdruckerei, V.; Winterthur, K. Multilingual Glossary of Terms Used In Wood Anatomy. *Comm. Nomencl. Int. Assoc. Wood Anat.* **1964**, 1–26.

13. Spicer, R. Symplasmic Networks in Secondary Vascular Tissues: Parenchyma Distribution and Activity Supporting Long-Distance Transport. *J. Exp. Bot.* **2014**, *65*, 1829–1848, doi:10.1093/jxb/ert459.

14. Taylor, A.M.; Gartner, B.L.; Morrell, J.J. Heartwood Formation and Natural Durability - A Review. *Wood Fiber Sci.* **2002**, *34*, 587–611.

15. Shain, L.; Mackay, J.F.G. Seasonal Fluctuation in Respiration of Aging Xylem in Relation to Heartwood Formation in Pinus Radiata . *Can. J. Bot.* **1973**, *51*, 737–741, doi:10.1139/b73-092.

16. Shain, L.; Hillis, W.E. Ethylene Production in Xylem of Pinus Radiata in Relation to Heartwood Formation . *Can. J. Bot.* **1973**, *51*, 1331–1335, doi:10.1139/b73-166.

17. Phillips, M.A.; Croteau, R.B. Resin-Based Defenses in Conifers. *Trends Plant Sci.* **1999**, *4*, 184–190, doi:10.1016/S1360-1385(99)01401-6.

18. Hudgins, J.W.; Franceschi, V.R. Methyl Jasmonate-Induced Ethylene Production Is Responsible for Conifer Phloem Defense Responses and Reprogramming of Stem Cambial Zone for Traumatic Resin Duct Formation. *Plant Physiol.* **2004**, *135*, 2134–2149, doi:10.1104/pp.103.037929.

19. Lim, K.J.; Paasela, T.; Harju, A.; Venäläinen, M.; Paulin, L.; Auvinen, P.; Kärkkäinen, K.; Teeri, T.H. Developmental Changes in Scots Pine Transcriptome during Heartwood Formation. *Plant Physiol.* **2016**, *172*, 1403–1417, doi:10.1104/pp.16.01082.

20. Magel, E.; Jay-Allemand, C.; Ziegler, H. Formation of Heartwood Substances in the Stemwood of Robinia Pseudoacacia L. II. Distribution of Nonstructural Carbohydrates and Wood Extractives across the Trunk. *Trees* **1994**, *8*, 165–171, doi:10.1007/BF00196843.

21. Bergström, B.; Gustafsson, G.; Gref, R.; Ericsson, A. Seasonal Changes of Pinosylvin Distribution in the Sapwood/Heartwood Boundary of Pinus Sylvestris. *Trees - Struct. Funct.* **1999**, *14*, 65–71, doi:10.1007/s004680050210.

22. Bergström, B. Chemical and Structural Changes during Heartwood Formation in Pinus Sylvestris. *Forestry* **2003**, *76*, 45–53, doi:10.1093/forestry/76.1.45.

23. Magel, E.; Abdel-Latif, A.; Hampp, R. Non-Structural Carbohydrates and Catalytic Activities of Sucrose Metabolizing Enzymes in Trunks of Two Juglans Species and Their Role in Heartwood Formation. *Holzforschung* **2001**, *55*, 135–145, doi:10.1515/HF.2001.022.

24. Magel, E.A. Biochemistry and Physiology of Heartwood Formation. *BIOS Sci. Publ. Oxford* **2000**, 363–376, doi:10.1136/bmj.2.4990.464.

25. Tsao, N.W.; Sun, Y.H.; Chien, S.C.; Chu, F.H.; Chang, S.T.; Kuo, Y.H.; Wang, S.Y. Content and Distribution of Lignans in Taiwania Cryptomerioides Hayata. *Holzforschung* **2016**, *70*, 511–518, doi:10.1515/hf-2015-0154.

26. Celedon, J.M.; Chiang, A.; Yuen, M.M.S.; Diaz-Chavez, M.L.; Madilao, L.L.; Finnegan, P.M.; Barbour, E.L.; Bohlmann, J. Heartwood-Specific Transcriptome and Metabolite Signatures of Tropical Sandalwood (Santalum Album) Reveal the Final Step of (Z)-Santalol Fragrance Biosynthesis. *Plant J.* **2016**, *86*, 289–299, doi:10.1111/tpj.13162.

27. Baldovini, N.; Delasalle, C.; Joulain, D. Phytochemistry of the Heartwood from Fragrant Santalum Species: A Review. *Flavour Fragr. J.* **2011**, *26*, 7–26, doi:10.1002/ffj.2025.

28. Nobuchi, T.; Kuroda, K.; Iwata, R.; Harada, H. Cytological Study of the Seasonal Features of Heartwood Formation of Sugi (Cryptomeria Japonica D. Don). *Mokuzai Gakkaishi* **1982**, *28*, 669–676.

29. Hennon, P.E. Are Heart Rot Fungi Major Factors of Disturbance in Gap-Dynamic Forests? *Northwest Sci.* **1995**, *69*, 284–293.

30. Van der Kamp, B.J. Effects of Heartwood Inhabiting Fungi on Thujaplicin Content and Decay Resistance of Western Redcedar (Thuja Plicata Donn.). *Wood Fiber Sci.* **1986**, *18*, 421–427.

31. Apolinário, F.E.; Martius, C. Ecological Role of Termites (Insecta, Isoptera) in Tree Trunks in Central Amazonian Rain Forests. *For. Ecol. Manage.* **2004**, *194*, 23–28, doi:10.1016/j.foreco.2004.01.052.

32. Buckland, D.C. Investigations of Decay in Western Red Cedar in British Columbia. *Can. J. Res.* **1946**, *24c*, 158–181, doi:10.1139/cjr46c-018.

33. McMurtrey, S.; Alcalá-Briseño, R.I.; Showalter, D.N.; LeBoldus, J.M. Draft Genome Resource for Forest Pathogen Coniferiporia Weirii - A Facultative Pathogen of Thuja Plicata and Callitropsis Nootkatensis. *Plant Dis.* **2023**, *107*, 534–537, doi:10.1094/PDIS-04-22-0917-A.

34. Allen, E.A.; Morrison, D.J.; Wallis, G.W. *Common Tree Diseases of British Columbia*; 1996; ISBN 0-662-24870-8.

35. Huang, Y.-T.; Eickwort, J.; Hulcr, J. Red Heart Disease in Pines Caused by Porodaedalea (Phellinus) Pini in Florida. *Edis* **2020**, *2020*, 3, doi:10.32473/edis-fr425-2019.

36. Allen, B.; Dalponte, M.; Hietala, A.M.; Ørka, H.O.; Næsset, E.; Gobakken, T. Detection of Root, Butt, and Stem Rot Presence in Norway Spruce with Hyperspectral Imagery. *Silva Fenn.* **2022**, *56*, 1–16, doi:10.14214/sf.10606.

37. Whitney, R.D. Root Rot of Spruce and Balsam Fir in Northwestern Ontario II. Causal Fungi and Site Relationships. **1979**, 42 pp.

38. Žemaitis, P.; Stakenas, V. Ecological Factors Influencing Frequency of Norway Spruce Butt Rot in Mature Stands in Lithuania. *Russ. J. Ecol.* **2016**, *47*, 355–363, doi:10.1134/S1067413616040172.

39. Hills, W.E. Heartwood and Tree Exudates. In *Springer-Verlag. Berlin, Germany*; 1987; p. 268 ISBN 9788578110796.

40. Kampe, A., Magel, E. New Insights into Heartwood and Heartwood Formation. In *Cellular Aspects of Wood Formation.*; 2013; Vol. 20, pp. 71–95 ISBN 9783642364907.

41. Stirling, R.; Sturrock, R.N.; Braybrooks, A. Fungal Decay of Western Redcedar Wood Products– a Review. *Int. Biodeterior. Biodegrad.* **2017**, *125*, 105–115, doi:10.1016/j.ibiod.2017.09.001.

42. Manter, D.K.; Kelsey, R.G.; Karchesy, J.J. Antimicrobial Activity of Extractable Conifer Heartwood Compounds toward Phytophthora Ramorum. *J. Chem. Ecol.* **2007**, *33*, 2133–2147, doi:10.1007/s10886-007-9368-0.

43. Arima, Y.; Nakai, Y.; Hayakawa, R.; Nishino, T. Antibacterial Effect of β-Thujaplicin on Staphylococci Isolated from Atopic Dermatitis: Relationship between Changes in the Number of Viable Bacterial Cells and Clinical Improvement in an Eczematous Lesion of Atopic Dermatitis. *J. Antimicrob. Chemother.* **2003**, *51*, 113–122, doi:10.1093/jac/dkg037.

44. Chiang, N.; Ma, L.; Lee, Y.; Tsao, N.; Yang, C.; Wang, S. The Gene Expression and Enzymatic Activity of Pinoresinol-Lariciresinol Reductase during Wood Formation in Taiwania Cryptomerioides Hayata. **2018**.

45. Rudloff, E. von; Nair, G. V. The Sesquiterpene Alcohols of the Heartwood of Thuja Occidentals L. *Can. J. Chem.* **1964**, *42*, 421–425, doi:10.1139/v64-059.

46. FANG, J.M, CHEN, Y.C, WANG, B.W and CHENG, Y.. Terpenes from Heartwood of Juniperus Chinensis. In *Phytochemistry*; 1996; Vol. 41, pp. 1361–1365.

47. Du, T.; Shupe, T.F.; Hse, C.Y. Antifungal Activities of Three Supercritical Fluid Extracted Cedar Oils. *Holzforschung* **2011**, *65*, 277–284, doi:10.1515/HF.2011.005.

48. Eller, F.J.; Clausen, C.A.; Green, F.; Taylor, S.L. Critical Fluid Extraction of Juniperus Virginiana L. and Bioactivity of Extracts against Subterranean Termites and Wood-Rot Fungi. *Ind. Crops Prod.* **2010**, *32*, 481–485, doi:10.1016/j.indcrop.2010.06.018.

49. Miyamoto, B.T.; Konkler, M.J.; Leavengood, S.; Morrell, J.J.; Sinha, A. Durability Assessment of Western Juniper from Five Different Growing Regions. *Wood Fiber Sci.* **2019**, *51*, 88–95, doi:10.22382/wfs-2019-009.

50. Harju, A.M.; Kainulainen, P.; Venäläinen, M.; Tiitta, M.; Viitanen, H. Differences in Resin Acid Concentration between Brown-Rot Resistant and Susceptible Scots Pine Heartwood. *Holzforschung* **2002**, *56*, 479–486, doi:10.1515/HF.2002.074.

51. Chen, S.Y.; Yen, P.L.; Chang, T.C.; Chang, S.T.; Huang, S.K.; Yeh, T.F. Distribution of Living Ray Parenchyma Cells and Major Bioactive Compounds During the Heartwood Formation of Taiwania Cryptomerioides Hayata. *J. Wood*

Chem. Technol. **2018**, *38*, 84–95, doi:10.1080/02773813.2017.1372478.

52. Wang, S.-C.; Hart, J.H. Heartwood Extractives of Maclura Pomifera and Their Role in Decay Resistance'. *Wood Fiber Sci.* **1981**, *15*, 290–301.

53. Dünisch, O.; Richter, H.G.; Koch, G. Wood Properties of Juvenile and Mature Heartwood in Robinia Pseudoacacia L. *Wood Sci. Technol.* **2010**, *44*, 301–313, doi:10.1007/s00226-009-0275-0.

54. Gupta, S. R., Ravindranah, B., and Seshari, T.R. Juglandaceae: Polyphenols of Juglans Nigra. *Phytochem.* **1972**, *11*, 2634–2636, doi:10.1139/b73-166.

55. McDaniel, C.A. Major Antitermitic Components of the Heartwood of Southern Catalpa. *J. Chem. Ecol.* **1992**, *18*, 359–369, doi:10.1007/BF00994237.

56. Rahman, A.A.; Samoylenko, V.; Jacob, M.R.; Sahu, R.; Jain, S.K.; Khan, S.I.; Tekwani, B.L.; Muhammad, I. Antiparasitic and Antimicrobial Indolizidines from the Leaves of Prosopis Glandulosa Var Glandulosa. *Planta Med.* **2011**, *77*, 1639–1643, doi:10.1055/s-0030-1270906.

57. Rajput, A.; Sharma, R.; Bharti, R. Pharmacological Activities and Toxicities of Alkaloids on Human Health. *Mater. Today Proc.* **2021**, *48*, 1407–1415, doi:10.1016/j.matpr.2021.09.189.

58. Kimmey, J.W. Cull Factors for Sitka Spruce, Western Hemlock, and Western Redcedar in Southeast Alaska. In *Station Pap. No. 6. USDA. Forest Service, Alaska Forest Research Center*; 1956; p. 31.

59. Freitag, C.M.; Morrell, J.J. Durability of Changing a Western Redcedar Resource. *Wood Fiber Sci.* **2001**, *33*, 69–75.

60. Debell, J.D.; Morrell, J.J.; Gartner, B.L. Within-Stem Variation in Tropolone Content and Decay Resistance of Second-Growth Western Redcedar. *For. Sci.* **1999**, *45*, 101–107.

61. Daniels, C.R.; Russell, J.H. Analysis of Western Redcedar (Thuja Plicata Donn) Heartwood Components by HPLC as a Possible Screening Tool for Trees with Enhanced Natural Durability. *J. Chromatogr. Sci.* **2007**, *45*, 281–285, doi:10.1093/chromsci/45.5.281.

62. Swan, E. P. ; Jiang, K.S. Formation of Heartwood Extractives in Western Red Cedar. *TAPPI* **1970**, *53*, 844–846.

63. Russell, J.H. and C.R.D. Variation in Western Redcedar Heartwood Extractives. In Proceedings of the A tale of two cedars – International symposium on western redcedar and yellow-cedar. Gen. Tech. Rep. PNW-GTR-828; 2010; Vol. 828, pp. 83–86.

64. War, A.R.; Paulraj, M.G.; Ahmad, T.; Buhroo, A.A.; Hussain, B.; Ignacimuthu, S.; Sharma, H.C. Mechanisms of Plant Defense against Insect Herbivores. *Plant Signal. Behav.* **2012**, *7*, 1306–1320.

65. Teshome, D.T.; Zharare, G.E.; Naidoo, S. The Threat of the Combined Effect of Biotic and Abiotic Stress Factors in Forestry Under a Changing Climate. *Front. Plant Sci.* **2020**, *11*, doi:10.3389/fpls.2020.601009.

66. Franceschi, V.R.; Krokene, P.; Christiansen, E.; Krekling, T. Anatomical and Chemical Defenses of Conifer Bark against Bark Beetles and Other Pests. *New Phytol.* **2005**, *167*, 353–376, doi:10.1111/j.1469-8137.2005.01436.x.

67. Mumm, R.; Hilker, M. Direct and Indirect Chemical Defence of Pine against Folivorous Insects. *Trends Plant Sci.* **2006**, *11*, 351–358, doi:10.1016/j.tplants.2006.05.007.

68. Howe, G.A.; Jander, G. Plant Immunity to Insect Herbivores. *Annu. Rev. Plant Biol.* **2008**, *59*, 41–66, doi:10.1146/annurev.arplant.59.032607.092825.

69. Hanley, M.E.; Lamont, B.B.; Fairbanks, M.M.; Rafferty, C.M. Plant Structural Traits and Their Role in Anti-Herbivore Defence. *Perspect. Plant Ecol. Evol. Syst.* **2007**, *8*, 157–178, doi:10.1016/j.ppees.2007.01.001.

70. Aljbory, Z.; Chen, M.S. Indirect Plant Defense against Insect Herbivores: A Review. *Insect Sci.* **2018**, *25*, 2–23, doi:10.1111/1744-7917.12436.

71. Verhage, A.; van Wees, S.C.M.; Pieterse, C.M.J. Plant Immunity: It's the Hormones Talking, but What Do They Say? *Plant Physiol.* **2010**, *154*, 536–540, doi:10.1104/pp.110.161570.

72. Wink, M. Plant Secondary Metabolites Modulate Insect Behavior-Steps toward Addiction? *Front. Physiol.* **2018**, *9*, 1–9, doi:10.3389/fphys.2018.00364.

73. Kawai, Y.; Ono, E.; Mizutani, M. Evolution and Diversity of the 2-Oxoglutarate-Dependent Dioxygenase Superfamily in Plants. *Plant J.* **2014**, *78*, 328–343, doi:10.1111/tpj.12479.

74. Divekar, P.A.; Narayana, S.; Divekar, B.A.; Kumar, R.; Gadratagi, B.G.; Ray, A.; Singh, A.K.; Rani, V.; Singh, V.; Singh, A.K.; et al. Plant Secondary Metabolites as Defense Tools against Herbivores for Sustainable Crop Protection. *Int. J. Mol. Sci.* **2022**, *23*, doi:10.3390/ijms23052690.

75. Lincoln Taiz, E.Z. Secondary Metabolites and Plant Defense. In *Plant physiology*; 2002; p. 283 ISBN 0878938230.

76. Abeer H. Ali, Mostafa Abdelrahman, M.A.E.-S. Alkaloid Role in Plant Defense Response to Growth and Stress. In *Bioactive Molecules in Plant Defense: Signaling in Growth and Stress*; 2019; p. ISBN : 978-3-030-27164-0 ISBN 9783030271657.

77. Brahmkshatriya, P. P., & Brahmkshatriya, P.S. Terpenes: Chemistry, Biological Role, and Therapeutic Applications. *Nat. Prod.* **2013**, 2665–2691, doi:10.1007/978-3-642-22144-6.

78. Boncan, D.A.T.; Tsang, S.S.K.; Li, C.; Lee, I.H.T.; Lam, H.M.; Chan, T.F.; Hui,

J.H.L. Terpenes and Terpenoids in Plants: Interactions with Environment and Insects. *Int. J. Mol. Sci.* **2020**, *21*, 1–19, doi:10.3390/ijms21197382.

79. Schilmiller, A.L.; Schauvinhold, I.; Larson, M.; Xu, R.; Charbonneau, A.L.; Schmidt, A.; Wilkerson, C.; Last, R.L.; Pichersky, E. Monoterpenes in the Glandular Trichomes of Tomato Are Synthesized from a Neryl Diphosphate Precursor Rather than Geranyl Diphosphate. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 10865–10870, doi:10.1073/pnas.0904113106.

80. Demissie, Z.A.; Erland, L.A.E.; Rheault, M.R.; Mahmoud, S.S. The Biosynthetic Origin of Irregular Monoterpenes in Lavandula: Isolation and Biochemical Characterization of a Novel Cis-Prenyl Diphosphate Synthase Gene, Lavandulyl Diphosphate Synthase. *J. Biol. Chem.* **2013**, *288*, 6333–6341, doi:10.1074/jbc.M112.431171.

81. Rivera, S.B.; Swedlund, B.D.; King, G.J.; Bell, R.N.; Hussey, C.E.; Shattuck-Eidens, D.M.; Wrobel, W.M.; Peiser, G.D.; Poulter, C.D. Chrysanthemyl Diphosphate Synthase: Isolation of the Gene and Characterization of the Recombinant Non-Head-to-Tail Monoterpene Synthase from Chrysanthemum Cinerariaefolium. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 4373–4378, doi:10.1073/pnas.071543598.

82. Tetali, S.D. Terpenes and Isoprenoids: A Wealth of Compounds for Global Use. *Planta* **2019**, *249*, 1–8, doi:10.1007/s00425-018-3056-x.

83. Nagel, R.; Schmidt, A.; Peters, R.J. Isoprenyl Diphosphate Synthases: The Chain Length Determining Step in Terpene Biosynthesis. *Planta* **2019**, *249*, 9–20, doi:10.1007/s00425-018-3052-1.

84. Nagegowda, D.A.; Gupta, P. Advances in Biosynthesis, Regulation, and Metabolic Engineering of Plant Specialized Terpenoids. *Plant Sci.* **2020**, *294*, 110457, doi:10.1016/j.plantsci.2020.110457.

85. Toffolatti, S.L.; Maddalena, G.; Passera, A.; Casati, P.; Bianco, P.A.; Quaglino, F. *Role of Terpenes in Plant Defense to Biotic Stress*; Elsevier Inc., 2021; ISBN 9780128229194.

86. Lewinsohn, E.; Gijzen, M.; Savage, T.J.; Croteau, R. Defense Mechanisms of Conifers: Relationship of Monoterpene Cyclase Activity to Anatomical Specialization and Oleoresin Monoterpene Content. *Plant Physiol.* **1991**, *96*, 38–43, doi:10.1104/pp.96.1.38.

87. He, X.; Jiang, Y.; Chen, S.; Chen, F.; Chen, F. Terpenoids and Their Possible Role in Defense Against a Fungal Pathogen Alternaria Tenuissima in Chrysanthemum Morifolium Cultivars. *J. Plant Growth Regul.* **2023**, *42*, 1144–1157, doi:10.1007/s00344-022-10619-z.

88. Romagni, J.G.; Allen, S.N.; Dayan, F.E. Allelopathic Effects of Volatile Cineoles on Two Weedy Plant Species. *J. Chem. Ecol.* **2000**, *26*, 303–313, doi:10.1023/A:1005414216848.

89. Dodson, C.H.; Dressler, R.L.; Hills, H.G.; Adams, R.M.; Williams, N.H. Biologically Active Compounds in Orchid Fragrances. *Science (80-.).* **1969**, *164*, 1243–1249, doi:10.1126/science.164.3885.1243.

90. Yoshioka, M.; Adachi, A.; Sato, Y.; Doke, N.; Kondo, T.; Yoshioka, H. RNAi of the Sesquiterpene Cyclase Gene for Phytoalexin Production Impairs Pre- and Post-Invasive Resistance to Potato Blight Pathogens. *Mol. Plant Pathol.* **2019**, *20*, 907–922, doi:10.1111/mpp.12802.

91. Song, N.; Ma, L.; Wang, W.; Sun, H.; Wang, L.; Baldwin, I.T.; Wu, J. Capsidiol, a Defensive Sesquiterpene Produced by Wild Tobacco in Response to Attack from the Fungal Pathogen Alternaria Alternata, Is Regulated by an ERF2-like Transcription Factor. *bioRxiv* **2019**, 573675.

92. Tian, X.; Ruan, J.; Huang, J.; Fang, X.; Mao, Y.; Wang, L.; Chen, X.; Yang, C. Gossypol: Phytoalexin of Cotton. *Sci. China Life Sci.* **2016**, *59*, 122–129, doi:10.1007/s11427-016-5003-z.

93. Prota, N.; Bouwmeester, H.J.; Jongsma, M.A. Comparative Antifeedant Activities of Polygodial and Pyrethrins against Whiteflies (Bemisia Tabaci) and Aphids (Myzus Persicae). *Pest Manag. Sci.* **2014**, *70*, 682–688, doi:10.1002/ps.3610.

94. Moesta, P.; West, C.A. Casbene Synthetase: Regulation of Phytoalexin Biosynthesis in Ricinus Communis L. Seedlings. Purification of Casbene Synthetase and Regulation of Its Biosynthesis during Elicitation. *Arch. Biochem. Biophys.* **1985**, *238*, 325–333, doi:10.1016/0003-9861(85)90171-7.

95. Schmelz, E.A.; Huffaker, A.; Sims, J.W.; Christensen, S.A.; Lu, X.; Okada, K.; Peters, R.J. Biosynthesis, Elicitation and Roles of Monocot Terpenoid Phytoalexins. *Plant J.* **2014**, *79*, 659–678, doi:10.1111/tpj.12436.

96. Keeling, C.I.; Bohlmann, J. Diterpene Resin Acids in Conifers. *Phytochemistry* **2006**, *67*, 2415–2423, doi:10.1016/j.phytochem.2006.08.019.

97. Jassbi, A.R.; Gase, K.; Hettenhausen, C.; Schmidt, A.; Baldwin, I.T. Silencing Geranylgeranyl Diphosphate Synthase in Nicotiana Attenuata Dramatically Impairs Resistance to Tobacco Hornworm. *Plant Physiol.* **2008**, *146*, 974–986, doi:10.1104/pp.107.108811.

98. Vaughan, M.M.; Wang, Q.; Webster, F.X.; Kiemle, D.; Hong, Y.J.; Tantillo, D.J.; Coates, R.M.; Wray, A.T.; Askew, W.; O'Donnell, C.; et al. Formation of the Unusual Semivolatile Diterpene Rhizathalene by the Arabidopsis Class I Terpene Synthase TPS08 in the Root Stele Is Involved in Defense against Belowground Herbivory. *Plant Cell* **2013**, *25*, 1108–1125, doi:10.1105/tpc.112.100057.

99. Tholl, D. Biosynthesis and Biological Functions of Terpenoids in Plants. In *In: Schrader, J., Bohlmann, J. (eds) Biotechnology of Isoprenoids. Advances in Biochemical Engineering/Biotechnology,*; 2015; Vol. 148, pp. 289–317 ISBN 9783319201061.

100. Lv, Z.Y.; Sun, W.J.; Jiang, R.; Chen, J.F.; Ying, X.; Zhang, L.; Chen, W.S.

Phytohormones Jasmonic Acid, Salicylic Acid, Gibberellins, and Abscisic Acid Are Key Mediators of Plant Secondary Metabolites. *World J. Tradit. Chinese Med.* **2021**, *7*, 307–325, doi:10.4103/wjtcm.wjtcm_20_21.

101. Hashimoto, H., Uragami, C., Cogdell, R.J. Carotenoids and Photosynthesis. In *In: Stange, C. (eds) Carotenoids in Nature. Subcellular Biochemistry*; 2016; Vol. 79, pp. 111–139 ISBN 978-3-319-39124-3.

102. Riedlmeier, M.; Ghirardo, A.; Wenig, M.; Knappe, C.; Koch, K.; Georgii, E.; Dey, S.; Parker, J.E.; Schnitzler, J.P.; Vlot, A.C. Monoterpenes Support Systemic Acquired Resistance within and between Plants. *Plant Cell* **2017**, *29*, 1440–1459, doi:10.1105/tpc.16.00898.

103. Liu, M.; Lu, S. Plastoquinone and Ubiquinone in Plants: Biosynthesis, Physiological Function and Metabolic Engineering. *Front. Plant Sci.* **2016**, *7*, 1–18, doi:10.3389/fpls.2016.01898.

104. Rogowska, A.; Szakiel, A. Enhancement of Phytosterol and Triterpenoid Production in Plant Hairy Root Cultures—Simultaneous Stimulation or Competition? *Plants* **2021**, *10*, doi:10.3390/plants10102028.

105. Pollastri, S.; Baccelli, I.; Loreto, F. Isoprene: An Antioxidant Itself or a Molecule with Multiple Regulatory Functions in Plants? *Antioxidants* **2021**, *10*, 1–13, doi:10.3390/antiox10050684.

106. Hayat, U.; Idrees Jilani, M.; Rehman, R.; Nadeem, F. A Review on Eucalyptus Globulus: A New Perspective in Therapeutics. *Ijcbs* **2015**, *8*, 85–91.

107. Li, L.; Ma, X.W.; Zhan, R.L.; Wu, H.X.; Yao, Q.S.; Xu, W.T.; Luo, C.; Zhou, Y.G.; Liang, Q.Z.; Wang, S.B. Profiling of Volatile Fragrant Components in a Mini-Core Collection of Mango Germplasms from Seven Countries. *PLoS One* **2017**, *12*, 1–14, doi:10.1371/journal.pone.0187487.

108. Masyita, A.; Mustika Sari, R.; Dwi Astuti, A.; Yasir, B.; Rahma Rumata, N.; Emran, T. Bin; Nainu, F.; Simal-Gandara, J. Terpenes and Terpenoids as Main Bioactive Compounds of Essential Oils, Their Roles in Human Health and Potential Application as Natural Food Preservatives. *Food Chem. X* **2022**, *13*, 100217, doi:10.1016/j.fochx.2022.100217.

109. Fiedor, J.; Burda, K. Potential Role of Carotenoids as Antioxidants in Human Health and Disease. *Nutrients* **2014**, *6*, 466–488, doi:10.3390/nu6020466.

110. Vickers, C.E.; Gershenzon, J.; Lerdau, M.T.; Loreto, F. A Unified Mechanism of Action for Volatile Isoprenoids in Plant Abiotic Stress. *Nat. Chem. Biol.* **2009**, *5*, 283–291, doi:10.1038/nchembio.158.

111. Kamatou, G.P.P.; Vermaak, I.; Viljoen, A.M.; Lawrence, B.M. Menthol: A Simple Monoterpene with Remarkable Biological Properties. *Phytochemistry* **2013**, *96*, 15–25, doi:10.1016/j.phytochem.2013.08.005.

112. Van Geldre, E.; Vergauwe, A.; Van Den Eeckhout, E. State of the Art of the

Production of the Antimalarial Compound Artemisinin in Plants. *Plant Mol. Biol.* **1997**, *33*, 199–209, doi:10.1023/A:1005716600612.

113. Wang, G.; Tang, W.; Bidigare, R.R. Bioactive Terpenoids As Pharmaceutical Agents. **2005**, 197–227.

114. Majeed, M.; Nagabhushanam, K.; Natarajan, S.; Vaidyanathan, P.; Karri, S.K.; Jose, J.A. Efficacy and Safety of 1% Forskolin Eye Drops in Open Angle Glaucoma - An Open Label Study. *Saudi J. Ophthalmol.* **2015**, *29*, 197–200, doi:10.1016/j.sjopt.2015.02.003.

115. Gao, W.; Hillwig, M.L.; Huango, L.; Cui, G.; Wang, X.; Kong, J.; Yang, B.; Peters, R.J. A Functional Genomics Approach to Tanshinone Biosynthesis Provides Stereochemical Insights. *Org. Lett.* **2009**, *11*, 5170–5173, doi:10.1021/ol902051v.

116. Bishayee, A.; Ahmed, S.; Brankov, N.; Perloff, M. Triterpenoids as Potential Agents for the Chemoprevention and Therapy of Breast Cancer. *Front. Biosci.* **2011**, *16*, 980–996, doi:10.2741/3730.

117. Tapiero, H.; Townsend, D.M.; Tew, K.D. The Role of Carotenoids in the Prevention of Human Pathologies. *Biomed. Pharmacother.* **2004**, *58*, 100–110, doi:10.1016/j.biopha.2003.12.006.

118. Zhao, J.; Zhao, J. Plant Troponoids: Chemistry, Biological Activity, and Biosynthesis. *Curr. Med. Chem.* **2007**, *14*, 2597–2621, doi:10.2174/092986707782023253.

119. Zielińska-Błajet, M.; Feder-Kubis, J. Monoterpenes and Their Derivatives—Recent Development in Biological and Medical Applications. *Int. J. Mol. Sci.* **2020**, *21*, 1–38, doi:10.3390/ijms21197078.

120. Zhao, C.; Yu, Z.; Teixeira da Silva, J.A.; He, C.; Wang, H.; Si, C.; Zhang, M.; Zeng, D.; Duan, J. Functional Characterization of a Dendrobium Officinale Geraniol Synthase Doges1 Involved in Floral Scent Formation. *Int. J. Mol. Sci.* **2020**, *21*, 1–15, doi:10.3390/ijms21197005.

121. Ninkuu, V.; Zhang, L.; Yan, J.; Fu, Z.; Yang, T.; Zeng, H. Biochemistry of Terpenes and Recent Advances in Plant Protection. *Int. J. Mol. Sci.* **2021**, *22*, doi:10.3390/ijms22115710.

122. Moniodis, J.; Jones, C.G.; Barbour, E.L.; Plummer, J.A.; Ghisalberti, E.L.; Bohlmann, J. The Transcriptome of Sesquiterpenoid Biosynthesis in Heartwood Xylem of Western Australian Sandalwood (Santalum Spicatum). *Phytochemistry* **2015**, *113*, 79–86, doi:10.1016/j.phytochem.2014.12.009.

123. Chen, Y.F.; Yang, C.H.; Chang, M.S.; Ciou, Y.P.; Huang, Y.C. Foam Properties and Detergent Abilities of the Saponins from Camellia Oleifera. *Int. J. Mol. Sci.* **2010**, *11*, 4417–4425, doi:10.3390/ijms11114417.

124. Hausch, B.J.; Lorjaroenphon, Y.; Cadwallader, K.R. Flavor Chemistry of Lemon-Lime Carbonated Beverages. *J. Agric. Food Chem.* **2015**, *63*, 112–119,

doi:10.1021/jf504852z.

125. Lois, L.M., Rodríguez-Concepción, M., Gallego, F., Campos, N. and Boronat, A. Carotenoid Biosynthesis during Tomato Fruit Development: Regulatory Role of 1-Deoxy-D-Xylulose 5-Phosphate Synthase. *Plant J.* **2000**, *22*, 503–513.

126. Tippmann, S.; Chen, Y.; Siewers, V.; Nielsen, J. From Flavors and Pharmaceuticals to Advanced Biofuels: Production of Isoprenoids in Saccharomyces Cerevisiae. *Biotechnol. J.* **2013**, *8*, 1435–1444, doi:10.1002/biot.201300028.

127. Lichtenthaler, H.K. The 1-Deoxy-D-Xylulose-5-Phosphate Pathway of Isoprenoid Biosynthesis in Plants. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **1999**, *50*, 47–65, doi:10.1146/annurev.arplant.50.1.47.

128. Lange, B.M.; Rujan, T.; Martin, W.; Croteau, R. Isoprenoid Biosynthesis: The Evolution of Two Ancient and Distinct Pathways across Genomes. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 13172–13177, doi:10.1073/pnas.240454797.

129. Bach, T.J.; Boronat, A.; Caelles, C.; Ferrer, A.; Weber, T.; Wettstein, A. Aspects Related to Mevalonate Biosynthesis in Plants. *Lipids* **1991**, *26*, 637–648, doi:10.1007/BF02536429.

130. Liao, P.; Wang, H.; Hemmerlin, A.; Nagegowda, D.A.; Bach, T.J.; Wang, M.; Chye, M.L. Past Achievements, Current Status and Future Perspectives of Studies on 3-Hydroxy-3-Methylglutaryl-CoA Synthase (HMGS) in the Mevalonate (MVA) Pathway. *Plant Cell Rep.* **2014**, *33*, 1005–1022, doi:10.1007/s00299-014-1592-9.

131. Miziorko, H.M. Enzymes of the Mevalonate Pathway of Isoprenoid Biosynthesis. *Arch. Biochem. Biophys.* **2011**, *505*, 131–143, doi:10.1016/j.abb.2010.09.028.

132. Wang, Q.; Quan, S.; Xiao, H. Towards Efficient Terpenoid Biosynthesis: Manipulating IPP and DMAPP Supply. *Bioresour. Bioprocess.* **2019**, *6*, doi:10.1186/s40643-019-0242-z.

133. Kuzuyama, T.; Shimizu, T.; Takahashi, S.; Seto, H. Fosmidomycin, a Specific Inhibitor of 1-Deoxy-D-Xylulose 5-Phosphate Reductoisomerase in the Nonmevalonate Pathway for Terpenoid Biosynthesis. *Tetrahedron Lett.* **1998**, *39*, 7913–7916, doi:10.1016/S0040-4039(98)01755-9.

134. Zulak, K.G.; Bohlmann, J. Terpenoid Biosynthesis and Specialized Vascular Cells of Conifer Defense. *J. Integr. Plant Biol.* **2010**, *52*, 86–97, doi:10.1111/j.1744-7909.2010.00910.x.

135. Keeling, C.I.; Bohlmann, J. Genes, Enzymes and Chemicals of Terpenoid Diversity in the Constitutive and Induced Defence of Conifers against Insects and Pathogens. *New Phytol.* **2006**, *170*, 657–675, doi:10.1111/j.1469-8137.2006.01716.x.

136. Karunanithi, P.S.; Zerbe, P. Terpene Synthases as Metabolic Gatekeepers in the

Evolution of Plant Terpenoid Chemical Diversity. *Front. Plant Sci.* **2019**, *10*, 1–23, doi:10.3389/fpls.2019.01166.

137. Zulak, K.G.; Bohlmann, J. Terpenoid Biosynthesis and Specialized Vascular Cells of Conifer Defense. *J. Integr. Plant Biol.* **2010**, *52*, 86–97, doi:10.1111/j.1744-7909.2010.00910.x.

138. McKay, S.A.B.; Hunter, W.L.; Godard, K.A.; Wang, S.X.; Martin, D.M.; Bohlmann, J.; Plant, A.L. Insect Attack and Wounding Induce Traumatic Resin Duct Development and Gene Expression of (-)-Pinene Synthase in Sitka Spruce. *Plant Physiol.* **2003**, *133*, 368–378, doi:10.1104/pp.103.022723.

139. BANNAN, M.W. Vertical Resin Ducts in the Secondary Wood of the Abietineae. *New Phytol.* **1936**, *35*, 11–46, doi:10.1111/j.1469-8137.1936.tb06864.x.

140. Spicer, R.; Holbrook, N.M. Parenchyma Cell Respiration and Survival in Secondary Xylem: Does Metabolic Activity Decline with Cell Age? *Plant, Cell Environ.* **2007**, *30*, 934–943, doi:10.1111/j.1365-3040.2007.01677.x.

141. Franceschi, V.R.; Krekling, T.; Christiansen, E. Application of Methyl Jasmonate on Picea Abies (Pinaceae) Stems Induces Defense-Related Responses in Phloem and Xylem. *Am. J. Bot.* **2002**, *89*, 578–586, doi:10.3732/ajb.89.4.578.

142. Martin, D.; Tholl, D.; Gershenzon, J.; Bohlmann, J. Methyl Jasmonate Induces Traumatic Resin Ducts, Terpenoid Resin Biosynthesis, and Terpenoid Accumulation in Developing Xylem of Norway Spruce Stems. *Plant Physiol.* **2002**, *129*, 1003–1018, doi:10.1104/pp.011001.

143. Spicer, R. Senescence in Secondary Xylem: Heartwood Formation as an Active Developmental Program. In; Physiological ecology: Vascular Transport in Plants, 2005; pp. 457–475.

144. Jiang, S.Y.; Jin, J.; Sarojam, R.; Ramachandran, S. A Comprehensive Survey on the Terpene Synthase Gene Family Provides New Insight into Its Evolutionary Patterns. *Genome Biol. Evol.* **2019**, *11*, 2078–2098, doi:10.1093/gbe/evz142.

145. Tholl, D. Terpene Synthases and the Regulation, Diversity and Biological Roles of Terpene Metabolism. *Curr. Opin. Plant Biol.* **2006**, *9*, 297–304, doi:10.1016/j.pbi.2006.03.014.

146. Keeling, C.I.; Weisshaar, S.; Ralph, S.G.; Jancsik, S.; Hamberger, B.; Dullat, H.K.; Bohlmann, J. Transcriptome Mining, Functional Characterization, and Phylogeny of a Large Terpene Synthase Gene Family in Spruce (Picea Spp.). *BMC Plant Biol.* **2011**, *11*, doi:10.1186/1471-2229-11-43.

147. Aubourg, S.; Lecharny, A.; Bohlmann, J. Genomic Analysis of the Terpenoid Synthase (AtTPS) Gene Family of Arabidopsis Thaliana. *Mol. Genet. Genomics* **2002**, *267*, 730–745, doi:10.1007/s00438-002-0709-y.

148. Yu, J.; Hu, S.; Wang, J.; Wong, G.K.S.; Li, S.; Liu, B.; Deng, Y.; Dai, L.; Zhou, Y.; Zhang, X.; et al. A Draft Sequence of the Rice Genome (Oryza Sativa L. Ssp.

Indica). *Science (80-.).* **2002**, *296*, 79–92, doi:10.1126/science.1068037.

149. Torr, P.; Putnam, N.; Ralph, S.; Rombauts, S.; Salamov, A.; Schein, J.; Sterck, L.; Aerts, A. The Genome of Black Cottonwood ,. *Science (80-.).* **2006**, *313*, 1596–1604.

150. Martin, D.M.; Fäldt, J.; Bohlmann, J. Functional Characterization of Nine Norway Spruce TPS Genes and Evolution of Gymnosperm Terpene Synthases of the TPS-d Subfamily. *Plant Physiol.* **2004**, *135*, 1908–1927, doi:10.1104/pp.104.042028.

151. Keeling, C.I.; Weisshaar, S.; Lin, R.P.C.; Bohlmann, J. Functional Plasticity of Paralogous Diterpene Synthases Involved in Conifer Defense. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 1085–1089, doi:10.1073/pnas.0709466105.

152. Shalev, T.J.; El-Dien, O.G.; Yuen, M.M.S.; Shengqiang, S.; Jackman, S.D.; Warren, R.L.; Coombe, L.; van der Merwe, L.; Stewart, A.; Boston, L.B.; et al. The Western Redcedar Genome Reveals Low Genetic Diversity in a Self-Compatible Conifer. *Genome Res.* **2022**, *32*, 1952–1964, doi:10.1101/gr.276358.121.

153. Bohlmann, J.; Phillips, M.; Ramachandiran, V.; Katoh, S.; Croteau, R. CDNA Cloning, Characterization, and Functional Expression of Four New Monoterpene Synthase Members of The. *Arch. Biochem. Biophys.* **1999**, *368*, 232–243.

154. Byun-McKay, A.; Godard, K.A.; Toudefallah, M.; Martin, D.M.; Alfaro, R.; King, J.; Bohlmann, J.; Plant, A.L. Wound-Induced Terpene Synthase Gene Expression in Sitka Spruce That Exhibit Resistance or Susceptibility to Attack by the White Pine Weevil. *Plant Physiol.* **2006**, *140*, 1009–1021, doi:10.1104/pp.105.071803.

155. Ee, S.F.; Mohamed-Hussein, Z.A.; Othman, R.; Shaharuddin, N.A.; Ismail, I.; Zainal, Z. Functional Characterization of Sesquiterpene Synthase from Polygonum Minus. *Sci. World J.* **2014**, *2014*, doi:10.1155/2014/840592.

156. Hall, D.E.; Zerbe, P.; Jancsik, S.; Quesada, A.L.; Dullat, H.; Madilao, L.L.; Yuen, M.; Bohlmann, J. Evolution of Conifer Diterpene Synthases: Diterpene Resin Acid Biosynthesis in Lodgepole Pine and Jack Pine Involves Monofunctional and Bifunctional Diterpene Synthases. *Plant Physiol.* **2013**, *161*, 600–616, doi:10.1104/pp.112.208546.

157. Wilderman, P.R.; Xu, M.; Jin, Y.; Coates, R.M.; Peters, R.J. Identification of Syn-Pimara-7,15-Diene Synthase Reveals Functional Clustering of Terpene Synthases Involved in Rice Phytoalexin/Allelochemical Biosynthesis. *Plant Physiol.* **2004**, *135*, 2098–2105, doi:10.1104/pp.104.045971.

158. Keeling, C.I.; Dullat, H.K.; Yuen, M.; Ralph, S.G.; Jancsik, S.; Bohlmann, J. Identification and Functional Characterization of Monofunctional Ent-Copalyl Diphosphate and Ent-Kaurene Synthases in White Spruce Reveal Different Patterns for Diterpene Synthase Evolution for Primary and Secondary Metabolism in Gymnosperms. *Plant Physiol.* **2010**, *152*, 1197–1208, doi:10.1104/pp.109.151456.

159. Bohlmann, J.; Meyer-Gauen, G.; Croteau, R. Plant Terpenoid Synthases: Molecular Biology and Phylogenetic Analysis. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95*, 4126–4133, doi:10.1073/pnas.95.8.4126.

160. Trapp, S.C.; Croteau, R.B. Genomic Organization of Plant Terpene Synthases and Molecular Evolutionary Implications. *Genetics* **2001**, *158*, 811–832, doi:10.1093/genetics/158.2.811.

161. Williams, D.C.; McGarvey, D.J.; Katahira, E.J.; Croteau, R. Truncation of Limonene Synthase Preprotein Provides a Fully Active “pseudomature” Form of This Monoterpene Cyclase and Reveals the Function of the Amino-Terminal Arginine Pair. *Biochemistry* **1998**, *37*, 12213–12220, doi:10.1021/bi980854k.

162. Whittington, D.A.; Wise, M.L.; Urbansky, M.; Coates, R.M.; Croteau, R.B.; Christianson, D.W. Bornyl Diphosphate Synthase: Structure and Strategy for Carbocation Manipulation by a Terpenoid Cyclase. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 15375–15380, doi:10.1073/pnas.232591099.

163. Christianson, D.. Structural Biology and Chemistry of the Terpenoid Cyclases. *Chem. Rev.* **2006**, *106*, 3412–3442.

164. Davis, E.M.; Croteau, R. Cyclization Enzymes in the Biosynthesis of Monoterpenes, Sesquiterpenes, and Diterpenes. **2000**, *209*, 53–95, doi:10.1007/3-540-48146-x_2.

165. Cane, D.E.; Xue, Q.; Fitzsimons, B.C. Trichodiene Synthase. Probing the Role of the Highly Conserved Aspartate-Rich Region by Site-Directed Mutagenesis. *Biochemistry* **1996**, *35*, 12369–12376, doi:10.1021/bi961344y.

166. Peters, R.J. Two Rings in Them All: The Labdane-Related Diterpenoids. *Nat. Prod. Rep.* **2010**, *27*, 1521–1530, doi:10.1039/c0np00019a.

167. Chen, F.; Tholl, D.; Bohlmann, J.; Pichersky, E. The Family of Terpene Synthases in Plants: A Mid-Size Family of Genes for Specialized Metabolism That Is Highly Diversified throughout the Kingdom. *Plant J.* **2011**, *66*, 212–229, doi:10.1111/j.1365-313X.2011.04520.x.

168. Zerbe, P.; Hamberger, B.; Yuen, M.M.S.; Chiang, A.; Sandhu, H.K.; Madilao, L.L.; Nguyen, A.; Hamberger, B.; Bach, S.S.; Bohlmann, J. Gene Discovery of Modular Diterpene Metabolism in Nonmodel Systems. *Plant Physiol.* **2013**, *162*, 1073–1091, doi:10.1104/pp.113.218347.

169. Prisic, S.; Xu, J.; Coates, R.M.; Peters, R.J. Probing the Role of the DXDD Motif in Class II Diterpene Cyclases. *ChemBioChem* **2007**, *8*, 869–874, doi:10.1002/cbic.200700045.

170. Sato, T.; Hoshino, T. Functional Analysis of the DXDDTA Motif in Squalene-Hopene Cyclase by Site-Directed Mutagenesis Experiments: Initiation Site of the Polycyclization Reaction and Stabilization Site of the Carbocation Intermediate of the Initially Cyclized A-Ring. *Biosci. Biotechnol. Biochem.* **1999**, *63*, 2189–2198, doi:10.1271/bbb.63.2189.

171. Peters, R.J.; Ravn, M.M.; Coates, R.M.; Croteau, R.B. Bifunctional Abietadiene Synthase: Free Diffusive Transfer of the (+)-Copalyl Diphosphate Intermediate between Two Distinct Active Sites. *J. Am. Chem. Soc.* **2001**, *123*, 8974–8978, doi:10.1021/ja010670k.

172. Gao, Y.; Honzatko, R.B.; Peters, R.J. Terpenoid Synthase Structures: A so Far Incomplete View of Complex Catalysis. *Nat. Prod. Rep.* **2012**, *29*, 1153–1175, doi:10.1039/c2np20059g.

173. Zhou, K.; Peters, R.J. Electrostatic Effects on (Di)Terpene Synthase Product Outcome. *Chem. Commun.* **2011**, *47*, 4074–4080, doi:10.1039/c0cc02960b.

174. Xu, M.; Wilderman, P.R.; Peters, R.J. Following Evolution's Lead to a Single Residue Switch for Diterpene Synthase Product Outcome. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 7387–7401, doi:10.1073/pnas.0611454104.

175. Criswell, J.; Potter, K.; Shephard, F.; Beale, M.H.; Peters, R.J. A Single Residue Change Leads to a Hydroxylated Product from the Class II Diterpene Cyclization Catalyzed by Abietadiene Synthase. *Org. Lett.* **2012**, *14*, 5828–5831, doi:10.1021/ol3026022.

176. Wilderman, P.R.; Peters, R.J. A Single Residue Switch Converts Abietadiene Synthase into a Pimaradiene Specific Cyclase. *J. Am. Chem. Soc.* **2007**, *129*, 15736–15737, doi:10.1021/ja074977g.

177. Keeling, C.I.; Madilao, L.L.; Zerbe, P.; Dullat, H.K.; Bohlmann, J. The Primary Diterpene Synthase Products of Picea Abies Levopimaradiene/ Abietadiene Synthase (PaLAS) Are Epimers of a Thermally Unstable Diterpenol. *J. Biol. Chem.* **2011**, *286*, 21145–21153, doi:10.1074/jbc.M111.245951.

178. Hamberger, B.; Ohnishi, T.; Hamberger, B.; Séguin, A.; Bohlmann, J. Evolution of Diterpene Metabolism: Sitka Spruce CYP720B4 Catalyzes Multiple Oxidations in Resin Acid Biosynthesis of Conifer Defense against Insects. *Plant Physiol.* **2011**, *157*, 1677–1695, doi:10.1104/pp.111.185843.

179. Steele, C.L.; Crock, J.; Bohlmann, J.; Croteau, R. Sesquiterpene Synthases from Grand Fir (Abies Grandis): Comparison of Constitutive and Wound-Induced Activities, and CDNA Isolation, Characterization, and Bacterial Expression of δ-Selinene Synthase and γ- Humulene Synthase. *J. Biol. Chem.* **1998**, *273*, 2078–2089, doi:10.1074/jbc.273.4.2078.

180. Phillips, M.A.; Wildung, M.R.; Williams, D.C.; Hyatt, D.C.; Croteau, R. CDNA Isolation, Functional Expression, and Characterization of (+)-α-Pinene Synthase and (-)-α-Pinene Synthase from Loblolly Pine (Pinus Taeda): Stereocontrol in Pinene Biosynthesis. *Arch. Biochem. Biophys.* **2003**, *411*, 267–276, doi:10.1016/S0003-9861(02)00746-4.

181. Hsieh, H.L.; Ma, L.T.; Wang, S.Y.; Chu, F.H. Cloning and Expression of a Sesquiterpene Synthase Gene from Taiwania Cryptomerioides. *Holzforschung* **2015**, *69*, 1041–1048, doi:10.1515/hf-2014-0279.

182. Foster, A.J.; Hall, D.E.; Mortimer, L.; Abercromby, S.; Gries, R.; Gries, G.; Bohlmann, J.; Russell, J.; Mattsson, J. Identification of Genes in Thuja Plicata Foliar Terpenoid Defenses. *Plant Physiol.* **2013**, *161*, 1993–2004, doi:10.1104/pp.112.206383.

183. Markus Lange, B.; Ahkami, A. Metabolic Engineering of Plant Monoterpenes, Sesquiterpenes and Diterpenes-Current Status and Future Opportunities. *Plant Biotechnol. J.* **2013**, *11*, 169–196, doi:10.1111/pbi.12022.

184. Ro, D.K.; Arimura, G.I.; Lau, S.Y.W.; Piers, E.; Bohlmann, J. Loblolly Pine Abietadienol/Abietadienal Oxidase PtAO (CYP720B1) Is a Multifunctional, Multisubstrate Cytochrome P450 Monooxygenase. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 8060–8065, doi:10.1073/pnas.0500825102.

185. McConkey, M.E.; Gershenzon, J.; Croteau, R.B. Developmental Regulation of Monoterpene Biosynthesis in the Glandular Trichomes of Peppermint. *Plant Physiol.* **2000**, *122*, 215–223, doi:10.1104/pp.122.1.215.

186. Bouwmeester, H.J.; Konings, M.C.J.M.; Gershenzon, J.; Karp, F.; Croteau, R. Cytochrome P-450 Dependent (+)-Limonene-6-Hydroxylation in Fruits of Caraway (Carum Carvi). *Phytochemistry* **1999**, *50*, 243–248, doi:10.1016/S0031-9422(98)00516-0.

187. Funk, C.; Croteau, R. Diterpenoid Resin Acid Biosynthesis in Conifers: Characterization of Two Cytochrome P450-Dependent Monooxygenases and an Aldehyde Dehydrogenase Involved in Abietic Acid Biosynthesis. *Arch. Biochem. Biophys.* 1994, *308*, 258–266.

188. Funk, C, Lewinsohn, E, Vogel, B.S, Steele, C.L, Croteau, R. Regulation of Oleoresinosis in Grand Fir (Abies Grandis) (Coordinate Induction of Monoterpene and Diterpene Cyclases and Two Cytochrome P450-Dependent Diterpenoid Hydroxylases by Stem Wounding). *Plant Physiol.* **1994**, *106*, 999–1005, doi:10.1104/pp.116.4.1497.

189. Bathe, U.; Tissier, A. Cytochrome P450 Enzymes: A Driving Force of Plant Diterpene Diversity. *Phytochemistry* **2019**, *161*, 149–162, doi:10.1016/j.phytochem.2018.12.003.

190. Geisler, K.; Jensen, N.B.; Yuen, M.M.S.; Madilao, L.; Bohlmann, J. Modularity of Conifer Diterpene Resin Acid Biosynthesis: P450 Enzymes of Different CYP720B Clades Use Alternative Substrates and Converge on the Same Products. *Plant Physiol.* **2016**, *171*, 152–164, doi:10.1104/pp.16.00180.

191. Diaz-Chavez, M.L.; Moniodis, J.; Madilao, L.L.; Jancsik, S.; Keeling, C.I.; Barbour, E.L.; Ghisalberti, E.L.; Plummer, J.A.; Jones, C.G.; Bohlmann, J. Biosynthesis of Sandalwood Oil: Santalum Album CYP76F Cytochromes P450 Produce Santalols and Bergamotol. *PLoS One* **2013**, *8*, doi:10.1371/journal.pone.0075053.

192. Gesell, A.; Blaukopf, M.; Madilao, L.; Yuen, M.M.S.; Withers, S.G.; Mattsson, J.; Russell, J.H.; Bohlmann, J. The Gymnosperm Cytochrome P450 CYP750B1 Catalyzes Stereospecific Monoterpene Hydroxylation of (+)-Sabinene in Thujone

Biosynthesis in Western Redcedar. *Plant Physiol.* **2015**, *168*, 94–106, doi:10.1104/pp.15.00315.

193. Rontein, D.; Onillon, S.; Herbette, G.; Lesot, A.; Werck-Reichhart, D.; Sallaud, C.; Tissier, A. CYP725A4 from Yew Catalyzes Complex Structural Rearrangement of Taxa-4(5),11(12)-Diene into the Cyclic Ether 5(12)-Oxa-3(11)-Cyclotaxane. *J. Biol. Chem.* **2008**, *283*, 6067–6075, doi:10.1074/jbc.M708950200.

194. Cankar, K.; Van Houwelingen, A.; Goedbloed, M.; Renirie, R.; De Jong, R.M.; Bouwmeester, H.; Bosch, D.; Sonke, T.; Beekwilder, J. Valencene Oxidase CYP706M1 from Alaska Cedar (Callitropsis Nootkatensis). *FEBS Lett.* **2014**, *588*, 1001–1007, doi:10.1016/j.febslet.2014.01.061.

195. Harada, T.; Harada, E.; Sakamoto, R.; Ashitani, T.; Fujita, K. Regio- and Substrate-Specific Oxidative Metabolism of Terpinolene by Cytochrome P450 Monooxygenases in <I>Cupressus Lusitanica</I> Cultured Cells. *Am. J. Plant Sci.* **2012**, *03*, 268–275, doi:10.4236/ajps.2012.32032.

196. Wray, G.A.; Hahn, M.W.; Abouheif, E.; Balhoff, J.P.; Pizer, M.; Rockman, M. V.; Romano, L.A. The Evolution of Transcriptional Regulation in Eukaryotes. *Mol. Biol. Evol.* **2003**, *20*, 1377–1419, doi:10.1093/molbev/msg140.

197. Schluttenhofer, C.; Yuan, L. Regulation of Specialized Metabolism by WRKY Transcription Factors. *Plant Physiol.* **2015**, *167*, 295–306, doi:10.1104/pp.114.251769.

198. Chen, M.; Yan, T.; Shen, Q.; Lu, X.; Pan, Q.; Huang, Y.; Tang, Y.; Fu, X.; Liu, M.; Jiang, W.; et al. GLANDULAR TRICHOME-SPECIFIC WRKY 1 Promotes Artemisinin Biosynthesis in Artemisia Annua. *New Phytol.* **2017**, *214*, 304–316, doi:10.1111/nph.14373.

199. Xu, Y.H.; Wang, J.W.; Wang, S.; Wang, J.Y.; Chen, X.Y. Characterization of GaWRKY1, a Cotton Transcription Factor That Regulates the Sesquiterpene Synthase Gene (+)-δ-Cadinene Synthase-A. *Plant Physiol.* **2004**, *135*, 507–515, doi:10.1104/pp.104.038612.

200. Ishihama, N.; Yamada, R.; Yoshioka, M.; Katou, S.; Yoshioka, H. Phosphorylation of the Nicotiana Benthamiana WRKY8 Transcription Factor by MAPK Functions in the Defense Response. *Plant Cell* **2011**, *23*, 1153–1170, doi:10.1105/tpc.110.081794.

201. Qiu, D.; Xiao, J.; Xie, W.; Liu, H.; Li, X.; Xiong, L.; Wang, S. Rice Gene Network Inferred from Expression Profiling of Plants Overexpressing OsWRKY13, a Positive Regulator of Disease Resistance. *Mol. Plant* **2008**, *1*, 538–551, doi:10.1093/mplant/ssn012.

202. Li, S.; Zhang, P.; Zhang, M.; Fu, C.; Yu, L. Functional Analysis of a WRKY Transcription Factor Involved in Transcriptional Activation of the DBAT Gene in Taxus Chinensis. *Plant Biol.* **2013**, *15*, 19–26, doi:10.1111/j.1438-8677.2012.00611.x.

203. Sun, Y.; Niu, Y.; Xu, J.; Li, Y.; Luo, H.; Zhu, Y.; Liu, M.; Wu, Q.; Song, J.; Sun, C.; et al. Discovery of WRKY Transcription Factors through Transcriptome Analysis and Characterization of a Novel Methyl Jasmonate-Inducible PqWRKY1 Gene from Panax Quinquefolius. *Plant Cell. Tissue Organ Cult.* **2013**, *114*, 269–277, doi:10.1007/s11240-013-0323-1.

204. Yu, Z.X.; Wang, L.J.; Zhao, B.; Shan, C.M.; Zhang, Y.H.; Chen, D.F.; Chen, X.Y. Progressive Regulation of Sesquiterpene Biosynthesis in Arabidopsis and Patchouli (Pogostemon Cablin) by the MIR156-Targeted SPL Transcription Factors. *Mol. Plant* **2015**, *8*, 98–110, doi:10.1016/j.molp.2014.11.002.

205. Estévez, J.M.; Cantero, A.; Reindl, A.; Reichler, S.; León, P. 1-Deoxy-D-Xylulose-5-Phosphate Synthase, a Limiting Enzyme for Plastidic Isoprenoid Biosynthesis in Plants. *J. Biol. Chem.* **2001**, *276*, 22901–22909, doi:10.1074/jbc.M100854200.

206. Wang, H.; Nagegowda, D.A.; Rawat, R.; Bouvier-Navé, P.; Guo, D.; Bach, T.J.; Chye, M.L. Overexpression of Brassica Juncea Wild-Type and Mutant HMG-CoA Synthase 1 in Arabidopsis up-Regulates Genes in Sterol Biosynthesis and Enhances Sterol Production and Stress Tolerance. *Plant Biotechnol. J.* **2012**, *10*, 31–42, doi:10.1111/j.1467-7652.2011.00631.x.

207. Zhou, F.; Pichersky, E. More Is Better: The Diversity of Terpene Metabolism in Plants. *Curr. Opin. Plant Biol.* **2020**, *55*, 1–10, doi:10.1016/j.pbi.2020.01.005.

208. Martin, D.; Tholl, D.; Gershenzon, J.; Bohlmann, J. Methyl Jasmonate Induces Traumatic Resin Ducts, Terpenoid Resin Biosynthesis, and Terpenoid Accumulation in Developing Xylem of Norway Spruce Stems. *Plant Physiol.* **2002**, *129*, 1003–1018, doi:10.1104/pp.011001.

209. Zhao, J.; Fujita, K.; Yamada, J.; Sakai, K. Improved β-Thujaplicin Production in Cupressus Lusitanica Suspension Cultures by Fungal Elicitor and Methyl Jasmonate. *Appl. Microbiol. Biotechnol.* **2001**, *55*, 301–305, doi:10.1007/s002530000555.

210. DeBell, J.D.; Morrell, J.J.; Gartner, B.L. Tropolone Content of Increment Cores as an Indicator of Decay Resistance in Western Redcedar. *Wood Fiber Sci.* **1997**, *29*, 364–369.

211. Stirling, R.; Morris, P.I. New Perspectives on the Role of Extractives in the Durability of Western Redcedar. *Proc. Can. Wood Preserv. Assoc.* **2011**, *32*, 12.

212. Nault, J. Radial Distribution of Thujaplicins in Old Growth and Second Growth Western Red Cedar (Thuja Plicata Donn). *Wood Sci. Technol.* **1988**, *22*, 73–80, doi:10.1007/BF00353230.

213. Liu, S.; Yamauchi, H. Hinokitiol, a Metal Chelator Derived from Natural Plants, Suppresses Cell Growth and Disrupts Androgen Receptor Signaling in Prostate Carcinoma Cell Lines. *Biochem. Biophys. Res. Commun.* **2006**, *351*, 26–32, doi:10.1016/j.bbrc.2006.09.166.

214. Akers, H.A.; Abrego, V.A.; Garland, E. Thujaplicins from Thuja Plicata as Iron

Transport Agents for Salmonella Typhimurium. *J. Bacteriol.* **1980**, *141*, 164–168, doi:10.1128/jb.141.1.164-168.1980.

215. Maclean, H.; Gardner, J.A.F. Analytical Method for Thujaplicins. *Anal. Chem.* **1956**, *28*, 509–512, doi:10.1021/ac50161a029.

216. Diouf, P.N.; Delbarre, N.; Perrin, D.; Gérardin, P.; Rapin, C.; Jacquot, J.P.; Gelhaye, E. Influence of Tropolone on Poria Placenta Wood Degradation. *Appl. Environ. Microbiol.* **2002**, *68*, 4377–4382, doi:10.1128/AEM.68.9.4377-4382.2002.

217. Taylor, A.M.; Gartner, B.L.; Morrell, J.J.; Tsunoda, K. Effects of Heartwood Extractive Fractions of Thuja Plicata and Chamaecyparis Nootkatensis on Wood Degradation by Termites or Fungi. *J. Wood Sci.* **2006**, *52*, 147–153, doi:10.1007/s10086-005-0743-6.

218. Johansson, C.I.; Saddler, J.N.; Beatson, R.P. Characterization of the Polyphenolics Related to the Colour of Western Red Cedar (Thuja Plicata Donn) Heartwood. *Holzforschung* **2000**, *54*, 246–254, doi:10.1515/HF.2000.042.

219. Roff, J.W.; Atkinson, J.M. Toxicity Tests of Water Soluble Phenolic Fraction (Thujaplicin-Free) of Western Red Cedar. *Can. J. Bot. 32*, 308–309.

220. Gardner, J.A.F.; Barton, G.M.; Maclean, H. The Polyoxyphenols of Western Red Cedar (Thuja Plicata Donn.): I. Isolation and Preliminary Characterization of Plicatic Acid. *Can. J. Chem.* **1959**, *37*, 1703–1709, doi:10.1139/v59-246.

221. Meng, F.; Liu, X.; Li, C.; Peng, X.; Wang, Q.; Xu, Q.; Sui, J.; Zhao, G.; Lin, J. Hinokitiol Inhibits Aspergillus Fumigatus by Interfering with the Cell Membrane and Cell Wall. *Front. Microbiol.* **2023**, *14*, doi:10.3389/fmicb.2023.1132042.

222. Fujita, K.; Yamaguchi, T.; Itose, R.; Sakai, K. Biosynthetic Pathway of β-Thujaplicin in the Cupressus Lusitanica Cell Culture. *J. Plant Physiol.* **2000**, *156*, 462–467, doi:10.1016/S0176-1617(00)80160-1.

223. Fujita, K.; Bunyu, Y.; Kuroda, K.; Ashitani, T.; Shigeto, J.; Tsutsumi, Y. A Novel Synthetic Pathway for Tropolone Ring Formation via the Olefin Monoterpene Intermediate Terpinolene in Cultured Cupressus Lusitanica Cells. *J. Plant Physiol.* **2014**, *171*, 610–614, doi:10.1016/j.jplph.2013.12.016.

224. Bentley, R. A Fresh Look at Natural Tropolonoids. *Nat. Prod. Rep.* **2008**, *25*, 118–138, doi:10.1039/b711474e.

225. Inamori Y, Sakagami Y, Morita Y, Shibata M, Sugiura M, Kumeda Y, Okabe T, Tsujibo H, I.N. Antifungal Activity of Hinokitiol-Related Compounds on Wood-Rotting Fungi and Their Insecticidal Activities. *Biol Pharm Bull* **2000**, *8*, 995–997.

226. Zhao, J.; Sakai, K. Multiple Signalling Pathways Mediate Fungal Elicitor-Induced β-Thujaplicin Biosynthesis in Cupressus Lusitanica Cell Cultures. *J. Exp. Bot.* **2003**, *54*, 647–656, doi:10.1093/jxb/erg062.

227. De Alwis, R.; Fujita, K.; Ashitani, T.; Kuroda, K. Volatile and Non-Volatile

Monoterpenes Produced by Elicitor-Stimulated Cupressus Lusitanica Cultured Cells. *J. Plant Physiol.* **2009**, *166*, 720–728, doi:10.1016/j.jplph.2008.09.009.

228. Yamaguchi, T.; Itose, R.; Sakai, K. Biosynthesis of Heartwood Tropolones I. Incorporation of Mevalonate and Acetate into ß-Thujaplicin (Hinokitiol) of Cupressus Lusitanica Cell Cultures. *J. Fac. Agric. Kyushu Univ.* **1997**, *42*, 131–138, doi:10.5109/24200.

229. Zhao, J.; Fujita, K.; Sakai, K. Production of β-Thujaplicin in Cupressus Lusitanica Suspension Cultures Fed with Organic Acids and Monoterpenes. *Biosci. Biotechnol. Biochem.* **2001**, *65*, 1027–1032, doi:10.1271/bbb.65.1027.

230. Matsunaga, Y.; Fujita, K.; Yamada, J.; Ashitani, T.; Sakai, K. Monoterpenes Produced by Cupressus Lusitanica Cultured Cells Including a Novel Monoterpene (1S, 2S, 6S)-(+)-1,6-Epoxy-4(8)-p-Menthen-2-Ol. *Nat. Prod. Res.* **2003**, *17*, 441–443, doi:10.1080/1478641031000111525.

231. Davison, J.; Al Fahad, A.; Cai, M.; Song, Z.; Yehia, S.Y.; Lazarus, C.M.; Bailey, A.M.; Simpson, T.J.; Cox, R.J. Genetic, Molecular, and Biochemical Basis of Fungal Tropolone Biosynthesis. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 7642–7647, doi:10.1073/pnas.1201469109.

232. Xin, M.; Bugg, T.D.H. Biomimetic Formation of 2-Tropolones by Dioxygenase-Catalysed Ring Expansion of Substituted 2,4-Cyclohexadienones. *ChemBioChem* **2010**, *11*, 272–276, doi:10.1002/cbic.200900631.

233. Woodhouse, R.N.; McDonald, E.; Ramage, R.; Battersby, A.R. Biosynthesis. Part 29.1 Colchicine: Studies on the Ring Expansion Step Focusing on the Fate of the Hydrogens at C-3 of Autumnaline. *J. Chem. Soc. - Perkin Trans. 1* **1998**, 2995–3001, doi:10.1039/a803852j.

234. Herbert, R.B. The Biosynthesis of Plant Alkaloids and Nitrogenous Microbial Metabolites. *Nat. Prod. Rep.* **2003**, *20*, 494–508, doi:10.1039/b006522f.

235. Cochrane, R.V.K.; Vederas, J.C. Highly Selective but Multifunctional Oxygenases in Secondary Metabolism. *Acc. Chem. Res.* **2014**, *47*, 3148–3161, doi:10.1021/ar500242c.

236. Hagel, J.M.; Facchini, P.J. Expanding the Roles for 2-Oxoglutarate-Dependent Oxygenases in Plant Metabolism. *Nat. Prod. Rep.* **2018**, *35*, 721–734, doi:10.1039/c7np00060j.

237. Hu, Z.; Ren, L.; Bu, J.; Liu, X.; Li, Q.; Guo, W.; Ma, Y.; Wang, J.; Chen, T.; Wang, L.; et al. Functional Characterization of a 2OGD Involved in Abietane-Type Diterpenoids Biosynthetic Pathway in Salvia Miltiorrhiza. *Front. Plant Sci.* **2022**, *13*, 1–12, doi:10.3389/fpls.2022.947674.

238. Islam, S.; Leissing, T.M.; Chowdhury, R.; Hopkinson, R.J.; Schofield, C.J. 2-Oxoglutarate-Dependent Oxygenases. **2018**.

239. Hewitson, K.S.; Granatino, N.; Welford, R.W.D.; McDonough, M.A.; Schofield,

C.J. Oxidation by 2-Oxoglutarate Oxygenases: Non-Haem Iron Systems in Catalysis and Signalling. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* **2005**, *363*, 807–828, doi:10.1098/rsta.2004.1540.

240. Wei, C.L.; Yang, Y.B.; Wang, W.C.; Liu, W.C.; Hsu, J.S.; Tsai, Y.C. Engineering Streptomyces Clavuligerus Deacetoxycephalosporin C Synthase for Optimal Ring Expansion Activity toward Penicillin G. *Appl. Environ. Microbiol.* **2003**, *69*, 2306–2312, doi:10.1128/AEM.69.4.2306-2312.2003.

241. Hausinger, R.P. *Fe(II)/α-Ketoglutarate-Dependent Hydroxylases and Related Enzymes*; 2004; Vol. 39; ISBN 1040923049044.

242. Morrison, D.J., Merler, H., Norris, D. Detection, Recognition and Management of Armillaria and Phellinus Root Diseases in the Southern Interior of British Columbia 1992, FRDA II. FRDA Rep. 179.

243. Crulckshank, M.G.; Morrison, D.J.; Punja, Z.K. Incidence of Armillaria Species in Precommercial Thinning Stumps and Spread of Armillaria Ostoyae to Adjacent Douglas-Fir Trees. *Can. J. For. Res.* **1997**, *27*, 481–490, doi:10.1139/x96-185.

244. Morrison, D.J.; Cruickshank, M.G.; Lalumière, A. Control of Laminated and Armillaria Root Diseases by Stump Removal and Tree Species Mixtures: Amount and Cause of Mortality and Impact on Yield after 40 Years. *For. Ecol. Manage.* **2014**, *319*, 75–98, doi:10.1016/j.foreco.2014.02.007.

245. Cleary, M.R.; Arhipova, N.; Morrison, D.J.; Thomsen, I.M.; Sturrock, R.N.; Vasaitis, R.; Gaitnieks, T.; Stenlid, J. Stump Removal to Control Root Disease in Canada and Scandinavia: A Synthesis of Results from Long-Term Trials. *For. Ecol. Manage.* **2013**, *290*, 5–14, doi:10.1016/j.foreco.2012.05.040.

246. Brunette, M.; Caurla, S. An Economic Comparison of Risk Handling Measures against Hylobius Abietis and Heterobasidion Annosum in the Landes de Gascogne Forest. *Ann. For. Sci.* **2016**, *73*, 777–787, doi:10.1007/s13595-016-0568-z.

247. Bogdanski, B.E.C.; Cruickshank, M.; Mario Di Lucca, C.; Becker, E. Stumping out Tree Root Disease – An Economic Analysis of Controlling Root Disease, Including Its Effects on Carbon Storage in Southern British Columbia. *For. Ecol. Manage.* **2018**, *409*, 129–147, doi:10.1016/j.foreco.2017.11.012.

248. Modi, D.; Simard, S.; Lavkulich, L.; Hamelin, R.C.; Grayston, S.J. Stump Removal and Tree Species Composition Promote a Bacterial Microbiome That May Be Beneficial in the Suppression of Root Disease. *FEMS Microbiol. Ecol.* **2021**, *97*, 1–13, doi:10.1093/femsec/fiaa213.

249. Taylor, A.M.; Gartner, B.L.; Morrell, J.J. Western Redcedar Extractives: Is There a Role for the Silviculturist? *For. Prod. J.* **2006**, *56*, 58–63.

250. Larjavaara, M.; Muller-Landau, H.C. Rethinking the Value of High Wood Density. *Funct. Ecol.* **2010**, *24*, 701–705, doi:10.1111/j.1365-2435.2010.01698.x.

251. Cruickshank, M.G. Climate and Site Factors Affecting Survival and Yield of Douglas-Fir in the Cedar-Hemlock Ecosystem of the Southern Interior of British Columbia. *Forestry* **2017**, *90*, 219–233, doi:10.1093/forestry/cpw040.

252. Fettig, C.J.; Reid, M.L.; Bentz, B.J.; Sevanto, S.; Spittlehouse, D.L.; Wang, T. Changing Climates, Changing Forests: A Western North American Perspective. *J. For.* **2013**, *111*, 214–228, doi:10.5849/jof.12-085.

253. Camarero, J.J.; Fernández-Pérez, L.; Kirdyanov, A. V.; Shestakova, T.A.; Knorre, A.A.; Kukarskih, V. V.; Voltas, J. Minimum Wood Density of Conifers Portrays Changes in Early Season Precipitation at Dry and Cold Eurasian Regions. *Trees - Struct. Funct.* **2017**, *31*, 1423–1437, doi:10.1007/s00468-017-1559-x.

254. Chave, J.; Coomes, D.; Jansen, S.; Lewis, S.L.; Swenson, N.G.; Zanne, A.E. Towards a Worldwide Wood Economics Spectrum. *Ecol. Lett.* **2009**, *12*, 351–366, doi:10.1111/j.1461-0248.2009.01285.x.

255. Anfodillo, T.; Olson, M.E. Tree Mortality: Testing the Link Between Drought, Embolism Vulnerability, and Xylem Conduit Diameter Remains a Priority. *Front. For. Glob. Chang.* **2021**, *4*, 1–7, doi:10.3389/ffgc.2021.704670.

256. Gryc, V.; Vavrčík, H.; Horn, K. Density of Juvenile and Mature Wood of Selected Coniferous Species. *J. For. Sci.* **2011**, *57*, 123–130, doi:10.17221/18/2010-jfs.

257. Jin, L.; Wilson, J.W.; Swan, E.P. Thujin, a Novel Lactone Isolated from the Discolored Heartwood of Thuja Plicata Donn. . *Can. J. Chem.* **1988**, *66*, 51–53, doi:10.1139/v88-007.

258. Lehong Jin, Bart J. Van Der Kamp, Jack Wilson, and E.P.S. Biodegradation of Thujaplicins in Living Western Red Cedar. *Can. J. For. Res. 18*, 784–788.

259. Gonzalez, J.S. Growth, Properties and Uses of Western Red Cedar. *Forintek Canada Corp. Spec. Publ. No. SP-37R* **2004**, 37.

260. Russell, J.H.; Yanchuk, A.D. Breeding for Growth Improvement and Resistance to Multiple Pests in Thuja Plicata. In Proceedings of the Proceedings of the 4th International Workshop on Genetics of Host-Parasite Interactions in Forestry; 2012; pp. 40–44.

261. Russell, J.H.; Ferguson, D.C. Preliminary Results from Five Generations of a Western Redcedar (Thuja Plicata) Selection Study with Self-Mating. *Tree Genet. Genomes* **2008**, *4*, 509–518, doi:10.1007/s11295-007-0127-8.

262. Aranda, P.S.; LaJoie, D.M.; Jorcyk, C.L. Bleach Gel. *Natl. Institutes Heal.* **2012**, *33*, 366–369, doi:10.1002/elps.201100335.Bleach.

263. Wiggins, D.; Mattsson, Ji. Transcriptomic Responses Governing Heartwood Formation and Osmotic Stress Responses in Western Redcedar (Thuja Plicata). **2022**, 47–78.

264. Bateman, A.; Martin, M.J.; O'Donovan, C.; Magrane, M.; Apweiler, R.; Alpi, E.;

Antunes, R.; Arganiska, J.; Bely, B.; Bingley, M.; et al. UniProt: A Hub for Protein Information. *Nucleic Acids Res.* **2015**, *43*, D204–D212, doi:10.1093/nar/gku989.

265. Procter, J.B.; Carstairs, G.M.; Soares, B.; Mourão, K.; Ofoegbu, T.C.; Barton, D.; Lui, L.; Menard, A.; Sherstnev, N.; Roldan-Martinez, D.; et al. Alignment of Biological Sequences with Jalview. *Methods Mol. Biol.* **2021**, *2231*, 203–224, doi:10.1007/978-1-0716-1036-7_13.

266. Bannai, H.; Tamada, Y.; Maruyama, O.; Nakai, K.; Miyano, S. Extensive Feature Detection of N-Terminal Protein Sorting Signals. *Bioinformatics* **2002**, *18*, 298–305, doi:10.1093/bioinformatics/18.2.298.

267. Emanuelsson, O.; Nielsen, H.; Brunak, S.; Von Heijne, G. Predicting Subcellular Localization of Proteins Based on Their N-Terminal Amino Acid Sequence. *J. Mol. Biol.* **2000**, *300*, 1005–1016, doi:10.1006/jmbi.2000.3903.

268. Chu, F.H.; Kuo, P.M.; Chen, Y.R.; Wang, S.Y. Cloning and Characterization of α-Pinene Synthase from Chamaecyparis Formosensis Matsum. *Holzforschung* **2009**, *63*, 69–74, doi:10.1515/HF.2009.019.

269. Chapter 14 Quantitative Analysis By Gas Chromatography Response Factors. Determination. Accuracy and Precision. *J. Chromatogr. Libr.* **1988**, *42*, 587–627, doi:10.1016/S0301-4770(08)70086-1.

270. Magori, N.; Fujita, T.; Kumamoto, E. Hinokitiol Inhibits Compound Action Potentials in the Frog Sciatic Nerve. *Eur. J. Pharmacol.* **2018**, *819*, 254–260, doi:10.1016/j.ejphar.2017.12.014.

271. Kwon, Y.; Kim, H.S.; Kim, H.W.; Lee, D.W.; Choi, Y.H. Antifungal Activities of β-Thujaplicin Originated in Chamaecyparis Obtusa. *J. Appl. Biol. Chem.* **2017**, *60*, 265–269, doi:10.3839/jabc.2017.042.

272. Shalev, T.J.; Gamal El-Dien, O.; Yuen, M.M.S.; van der Merwe, L.; Kirst, M.; Yanchuk, A.D.; Ritland, C.; Russell, J.H.; Bohlmann, J. Genetic Architecture of Terpene Chemistry and Growth Traits and the Impact of Inbreeding on These Traits in Western Redcedar (Thuja Plicata). *Evol. Appl.* **2023**, *16*, 673–687, doi:10.1111/eva.13526.

273. Hoelscher, D.J.; Williams, D.C.; Wildung, M.R.; Croteau, R. A CDNA Clone for 3-Carene Synthase from Salvia Stenophylla. *Phytochemistry* **2003**, *62*, 1081–1086, doi:10.1016/S0031-9422(02)00674-X.

274. Huber, D.P.W.; Philippe, R.N.; Godard, K.A.; Sturrock, R.N.; Bohlmann, J. Characterization of Four Terpene Synthase CDNAs from Methyl Jasmonate-Induced Douglas-Fir, Pseudotsuga Menziesii. *Phytochemistry* **2005**, *66*, 1427–1439, doi:10.1016/j.phytochem.2005.04.030.

275. Adal, A.M.; Sarker, L.S.; Lemke, A.D.; Mahmoud, S.S. Isolation and Functional Characterization of a Methyl Jasmonate-Responsive 3-Carene Synthase from Lavandula x Intermedia. *Plant Mol. Biol.* **2017**, *93*, 641–657, doi:10.1007/s11103-017-0588-6.

276. Zulak, K.G.; Lippert, D.N.; Kuzyk, M.A.; Domanski, D.; Chou, T.; Borchers, C.H.; Bohlmann, J. Targeted Proteomics Using Selected Reaction Monitoring Reveals the Induction of Specific Terpene Synthases in a Multi-Level Study of Methyl Jasmonate-Treated Norway Spruce (Picea Abies). *Plant J.* **2009**, *60*, 1015–1030, doi:10.1111/j.1365-313X.2009.04020.x.

277. Dwivedi, V.; Rao, S.; Bomzan, D.P.; Kumar, S.R.; Shanmugam, P. V.; Olsson, S.B.; Nagegowda, D.A. Functional Characterization of a Defense-Responsive Bulnesol/Elemol Synthase from Potato. *Physiol. Plant.* **2021**, *171*, 7–21, doi:10.1111/ppl.13199.

278. Hall, D.E.; Robert, J.A.; Keeling, C.I.; Domanski, D.; Quesada, A.L.; Jancsik, S.; Kuzyk, M.A.; Hamberger, B.; Borchers, C.H.; Bohlmann, J. An Integrated Genomic, Proteomic and Biochemical Analysis of (+)-3-Carene Biosynthesis in Sitka Spruce (Picea Sitchensis) Genotypes That Are Resistant or Susceptible to White Pine Weevil. *Plant J.* **2011**, *65*, 936–948, doi:10.1111/j.1365-313X.2010.04478.x.

279. Roach, C.R.; Hall, D.E.; Zerbe, P.; Bohlmann, J. Plasticity and Evolution of (+)-3-Carene Synthase and (-)-Sabinene Synthase Functions of a Sitka Spruce Monoterpene Synthase Gene Family Associated with Weevil Resistance. *J. Biol. Chem.* **2014**, *289*, 23859–23869, doi:10.1074/jbc.M114.571703.

280. Robert, J.A.; Madilao, L.L.; White, R.; Yanchuk, A.; King, J.; Bohlmann, J. Terpenoid Metabolite Profiling in Sitka Spruce Identifies Association of Dehydroabietic Acid, (+)-3-Carene, and Terpinolene with Resistance against White Pine Weevil. *Botany* **2010**, *88*, 810–820, doi:10.1139/B10-049.

281. Liang, J.Y.; Guo, S.S.; Zhang, W.J.; Geng, Z.F.; Deng, Z.W.; Du, S.S.; Zhang, J. Fumigant and Repellent Activities of Essential Oil Extracted from Artemisia Dubia and Its Main Compounds against Two Stored Product Pests. *Nat. Prod. Res.* **2018**, *32*, 1234–1238, doi:10.1080/14786419.2017.1331227.

282. Cheng, S.S.; Lin, C.Y.; Chung, M.J.; Chang, S.T. Chemical Composition and Antitermitic Activity against Coptotermes Formosanus SHIRAKI of Cryptomeria Japonica Leaf Essential Oil. *Chem. Biodivers.* **2012**, *9*, 352–358, doi:10.1002/cbdv.201100243.

283. SeonHong, K.; SuYeon, L.; ChangYoung, H.; SeongMin, C.; MiJin, P.; InGyu, C. Antifungal Effect of Elemol and Eudesmol from Cryptomeria Japonica Essential Oil against Trichophyton Rubrum. *Acad. J. Agric. Res.* **2016**, *4*, 511–517, doi:10.15413/ajar.2016.0182.

284. Li, M.Z.; Elledge, S.J. Harnessing Homologous Recombination in Vitro to Generate Recombinant DNA via SLIC. *Nat. Methods* **2007**, *4*, 251–256, doi:10.1038/nmeth1010.

285. Cyr, A.; Wilderman, P.R.; Determan, M.; Peters, R.J. A Modular Approach for Facile Biosynthesis of Labdane-Related Diterpenes. *J. Am. Chem. Soc.* **2007**, *129*, 6684–6685, doi:10.1021/ja071158n.

286. Morrone, D.; Lowry, L.; Determan, M.K.; Hershey, D.M.; Xu, M.; Peters, R.J. Increasing Diterpene Yield with a Modular Metabolic Engineering System in E. Coli: Comparison of MEV and MEP Isoprenoid Precursor Pathway Engineering. *Appl. Microbiol. Biotechnol.* **2010**, *85*, 1893–1906, doi:10.1007/s00253-009-2219-x.

287. Reiling, K.K.; Yoshikuni, Y.; Martin, V.J.J.; Newman, J.; Bohlmann, J.; Keasling, J.D. Mono and Diterpene Production in Escherichia Coli. *Biotechnol. Bioeng.* **2004**, *87*, 200–212, doi:10.1002/bit.20128.

288. Xu, M.; Hillwig, M.L.; Tiernan, M.S.; Peters, R.J. Probing Labdane-Related Diterpenoid Biosynthesis in the Fungal Genus Aspergillus. *J. Nat. Prod.* **2017**, *80*, 328–333, doi:10.1021/acs.jnatprod.6b00764.

289. Martin, D.M.; Fäldt, J.; Bohlmann, J.; Martin, D.M.; Faldt, J.; Bohlmann, J. Functional Characterization of Nine Norway Spruce TPS Genes and Evolution of Gymnosperm Terpene Synthases of the TPS-d Subfamily Published by : American Society of Plant Biologists (ASPB) Stable URL : Https://Www.Jstor.Org/Stable/4356548 Linked Referenc. **2019**, *135*, 1908–1927, doi:10.1104/pp.104.042028.carbon.

290. Peters, R.J.; Flory, J.E.; Jetter, R.; Ravn, M.M.; Lee, H.J.; Coates, R.M.; Croteau, R.B. Abietadiene Synthase from Grand Fir (Abies Grandis): Characterization and Mechanism of Action of the “pseudomature” Recombinant Enzyme. *Biochemistry* **2000**, *39*, 15592–15602, doi:10.1021/bi001997l.

291. Peters, R.J.; Croteau, R.B. Abietadiene Synthase Catalysis: Conserved Residues Involved in Protonation-Initiated Cyclization of Geranylgeranyl Diphosphate to (+)-Copalyl Diphosphate. *Biochemistry* **2002**, *41*, 1836–1842, doi:10.1021/bi011879d.

292. Wu, Y.; Zhou, K.; Toyomasu, T.; Sugawara, C.; Oku, M.; Abe, S.; Usui, M.; Mitsuhashi, W.; Chono, M.; Chandler, P.M.; et al. Functional Characterization of Wheat Copalyl Diphosphate Synthases Sheds Light on the Early Evolution of Labdane-Related Diterpenoid Metabolism in the Cereals. *Phytochemistry* **2012**, *84*, 40–46, doi:10.1016/j.phytochem.2012.08.022.

293. Ma, L.T.; Lee, Y.R.; Tsao, N.W.; Wang, S.Y.; Zerbe, P.; Chu, F.H. Biochemical Characterization of Diterpene Synthases of Taiwania Cryptomerioides Expands the Known Functional Space of Specialized Diterpene Metabolism in Gymnosperms. *Plant J.* **2019**, 1254–1272, doi:10.1111/tpj.14513.

294. Murphy, K.M.; Ma, L.T.; Ding, Y.; Schmelz, E.A.; Zerbe, P. Functional Characterization of Two Class Ii Diterpene Synthases Indicates Additional Specialized Diterpenoid Pathways in Maize (Zea Mays). *Front. Plant Sci.* **2018**, *871*, 1–12, doi:10.3389/fpls.2018.01542.

295. Vogel, B.S.; Wildung, M.R.; Vogel, G.; Croteau, R. Abietadiene Synthase from Grand Fir (Abies Grandis). *J. Biol. Chem.* **1996**, *271*, 23262–23268, doi:10.1074/jbc.271.38.23262.

296. Schepmann, H.G.; Pang, J.; Matsuda, S.P.T. Cloning and Characterization of

Ginkgo Biloba Levopimaradiene Synthase, Which Catalyzes the First Committed Step in Ginkgolide Biosynthesis. *Arch. Biochem. Biophys.* **2001**, *392*, 263–269, doi:10.1006/abbi.2001.2438.

297. Ro, D.K.; Bohlmann, J. Diterpene Resin Acid Biosynthesis in Loblolly Pine (Pinus Taeda): Functional Characterization of Abietadiene/Levopimaradiene Synthase (PtTPS-LAS) CDNA and Subcellular Targeting of PtTPS-LAS and Abietadienol/Abietadienal Oxidase (PtAO, CYP720B1). *Phytochemistry* **2006**, *67*, 1572–1578, doi:10.1016/j.phytochem.2006.01.011.

298. Morrone, D.; Hillwig, M.L.; Mead, M.E.; Lowry, L.; Fulton, D.B.; Peters, R.J. Evident and Latent Plasticity across the Rice Diterpene Synthase Family with Potential Implications for the Evolution of Diterpenoid Metabolism in the Cereals. *Biochem. J.* **2011**, *435*, 589–595, doi:10.1042/BJ20101429.

299. Morrone, D.; Jin, Y.; Xu, M.; Choi, S.Y.; Coates, R.M.; Peters, R.J. An Unexpected Diterpene Cyclase from Rice: Functional Identification of a Stemodene Synthase. *Arch. Biochem. Biophys.* **2006**, *448*, 133–140, doi:10.1016/j.abb.2005.09.001.

300. Stirling, R.; Morris, P.I. Potential Contributions of Lignans to Decay Resistance in Western Red Cedar. *Wood Sci. Technol.* **2016**, *50*, 399–412, doi:10.1007/s00226-015-0784-y.

301. Schultz, T.P.; Nicholas, D.D. Naturally Durable Heartwood: Evidence for a Proposed Dual Defensive Function of the Extractives. *Phytochemistry* **2000**, *54*, 47–52, doi:10.1016/S0031-9422(99)00622-6.

302. ERDTMAN, H.; GRIPENBERG, J. Antibiotic Substances from the Heart Wood of Thuja Plicata Don. *Nature* **1948**, *161*.

303. Iwamoto, M.; Ohtsu, H.; Tokuda, H.; Nishino, H.; Matsunaga, S.; Tanaka, R. Anti-Tumor Promoting Diterpenes from the Stem Bark of Thuja Standishii (Cupressaceae). *Bioorganic Med. Chem.* **2001**, *9*, 1911–1921, doi:10.1016/S0968-0896(01)00099-2.

304. Azémard, C.; Ménager, M.; Vieillescazes, C. On the Tracks of Sandarac, Review and Chemical Analysis. *Environ. Sci. Pollut. Res.* **2017**, *24*, 27746–27754, doi:10.1007/s11356-017-0522-0.

305. Sugimoto, N.; Kuroyanagi, M.; Kato, T.; Sato, K.; Tada, A.; Yamazaki, T.; Tanamoto, K. Identification of the Main Constituents in Sandarac Resin, a Natural Gum Base. *J. Food Hyg. Soc. Japan* **2006**, *47*, 76–79, doi:10.3358/shokueishi.47.76.

306. Edwards, O.E.; Nicolson, A.; Rodger, M.N. The Structure of Sandaracopimaric Acid. *Can. J. Chem.* **1960**, *38*, 663–667, doi:10.1139/v60-096.

307. Simoneit, B.R.T.; Cox, R.E.; Oros, D.R.; Otto, A. Terpenoid Compositions of Resins from Callitris Species (Cupressaceae). *Molecules* **2018**, *23*, 1–13, doi:10.3390/molecules23123384.

308. Micales, J.A.; Han, J.S.; Davis, J.L.; Young, R.A. Chemical Composition and Fungitoxic Activities of Pine Cone Extractives. *Mycotoxins, Wood Decay, Plant Stress. Biocorrosion, Gen. Biodeterior.* **1994**, *4*, 317–332, doi:10.1007/978-1-4757-9450-2_25.

309. Ejike, C.E.C.C.; Gong, M.; Udenigwe, C.C. Phytoalexins from the Poaceae: Biosynthesis, Function and Prospects in Food Preservation. *Food Res. Int.* **2013**, *52*, 167–177, doi:10.1016/j.foodres.2013.03.012.

310. Peters, R.J. Uncovering the Complex Metabolic Network Underlying Diterpenoid Phytoalexin Biosynthesis in Rice and Other Cereal Crop Plants. *Phytochemistry* **2006**, *67*, 2307–2317, doi:10.1016/j.phytochem.2006.08.009.

311. Koga, J.; Shimura, M.; Oshima, K.; Ogawa, N.; Yamauchi, T.; Ogasawara, N. Phytocassanes A, B, C and D, Novel Diterpene Phytoalexins from Rice, Oryza Sativa L. *Tetrahedron* **1995**, *51*, 7907–7918, doi:10.1016/0040-4020(95)00423-6.

312. Koga, J.; Ogawa, N.; Yamauchi, T.; Klkuchi, M.; Ogasawara, N.; Shimura, M. Functional Moiety for the Antifungal Activity of Phytocassane E, a Diterpene Phytoalexin from Rice. *Phytochemistry* **1997**, *44*, 249–253, doi:10.1016/S0031-9422(96)00534-1.

313. Fukuta, M.; Xuan, T.D.; Deba, F.; Tawata, S.; Khanh, T.D.; Chung, I.M. Comparative Efficacies in Vitro of Antibacterial, Fungicidal, Antioxidant, and Herbicidal Activities of Momilatones A and B. *J. Plant Interact.* **2007**, *2*, 245–251, doi:10.1080/17429140701713811.

314. Tamura, K.; Stecher, G.; Kumar, S. MEGA11: Molecular Evolutionary Genetics Analysis Version 11. *Mol. Biol. Evol.* **2021**, *38*, 3022–3027, doi:10.1093/molbev/msab120.

315. Nicholas, K.B., & Nicholas, H.B. GeneDoc: A Tool for Editing and Annotating Multiple Sequence Alignments. **1997**.

316. Letunic, I.; Bork, P. Interactive Tree of Life (ITOL) v5: An Online Tool for Phylogenetic Tree Display and Annotation. *Nucleic Acids Res.* **2021**, *49*, W293–W296, doi:10.1093/nar/gkab301.

317. Norkunas, K.; Harding, R.; Dale, J.; Dugdale, B. Improving Agroinfiltration-Based Transient Gene Expression in Nicotiana Benthamiana. *Plant Methods* **2018**, *14*, 1–14, doi:10.1186/s13007-018-0343-2.

318. Doyon, T.J.; Skinner, K.C.; Yang, D.; Mallik, L.; Wymore, T.; Koutmos, M.; Zimmerman, P.M.; Narayan, A.R.H. Radical Tropolone Biosynthesis. *ChemRxiv* **2020**, 11–13.

319. Cox, R.J.; Al-Fahad, A. Chemical Mechanisms Involved during the Biosynthesis of Tropolones. *Curr. Opin. Chem. Biol.* **2013**, *17*, 532–536, doi:10.1016/j.cbpa.2013.06.029.

320. Farrow, S.C.; Facchini, P.J. Functional Diversity of 2-Oxoglutarate/Fe(II)-

Dependent Dioxygenases in Plant Metabolism. *Front. Plant Sci.* **2014**, *5*, 1–15, doi:10.3389/fpls.2014.00524.

321. Martinez, S.; Hausinger, R.P. Catalytic Mechanisms of Fe(II)- and 2-Oxoglutarate-Dependent Oxygenases. *J. Biol. Chem.* **2015**, *290*, 20702–20711, doi:10.1074/jbc.R115.648691.

322. Itoh, H.; Ueguchi-Tanaka, M.; Sentoku, N.; Kitano, H.; Matsuoka, M.; Kobayashi, M. Cloning and Functional Analysis of Two Gibberellin 3β-Hydroxylase Genes That Are Differently Expressed during the Growth of Rice. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 8909–8914, doi:10.1073/pnas.141239398.

323. Jiang, X.; Shi, Y.; Fu, Z.; Li, W.W.; Lai, S.; Wu, Y.; Wang, Y.; Liu, Y.; Gao, L.; Xia, T. Functional Characterization of Three Flavonol Synthase Genes from Camellia Sinensis: Roles in Flavonol Accumulation. *Plant Sci.* **2020**, *300*, 110632, doi:10.1016/j.plantsci.2020.110632.

324. Zhao, J.; Matsunaga, Y.; Fujita, K.; Sakai, K. Signal Transduction and Metabolic Flux of β-Thujaplicin and Monoterpene Biosynthesis in Elicited Cupressus Lusitanica Cell Cultures. *Metab. Eng.* **2006**, *8*, 14–29, doi:10.1016/j.ymben.2005.09.002.

325. Tasnim, S.; Gries, R.; Mattsson, J. Identification of Three Monofunctional Diterpene Synthases with Specific Enzyme Activities Expressed during Heartwood Formation in Western Redcedar (Thuja Plicata) Trees. *Plants* **2020**, *9*, 1–18, doi:10.3390/plants9081018.

326. Li, B.; Dewey, C.N. RSEM: Accurate Transcript Quantification from RNA-Seq Data with or without a Reference Genome. *BMC Bioinformatics* **2011**, *12*, doi:10.1186/1471-2105-12-323.

327. Ritchie, M.E.; Phipson, B.; Wu, D.; Hu, Y.; Law, C.W.; Shi, W.; Smyth, G.K. Limma Powers Differential Expression Analyses for RNA-Sequencing and Microarray Studies. *Nucleic Acids Res.* **2015**, *43*, e47, doi:10.1093/nar/gkv007.

328. Robinson, M.D.; Oshlack, A. A Scaling Normalization Method for Differential Expression Analysis of RNA-Seq Data. *Genome Biol.* **2010**, *11*, 1–9.

329. Law, C.W.; Smyth, G.K.; Ritchie, M.E.; Zeglinski, K.; Dong, X.; Alhamdoosh, M. A Guide to Creating Design Matrices for Gene Expression Experiments. *F1000Research* **2020**, *9*, 1–45, doi:10.12688/f1000research.27893.1.

330. Law, C.W.; Chen, Y.; Shi, W.; Smyth, G.K. Voom: Precision Weights Unlock Linear Model Analysis Tools for RNA-Seq Read Counts. *Genome Biol.* **2014**, *15*, 1–17, doi:10.1186/gb-2014-15-2-r29.

331. Young, M.D.; Wakefield, M.J.; Smyth, G.K.; Oshlack, A. Gene Ontology Analysis for RNA-Seq: Accounting for Selection Bias. *Genome Biol.* **2010**, *11*, doi:10.1186/gb-2010-11-2-r14.

332. Morrison, D.J.; Wallis, G.W.; Weir, L.C. Control of Armillaria and Phellinus Root

Diseases: 20-Year Results from the Skimikin Stump Removal Experiment. *Can. For. Serv. Pacific For. Cent.* **1988**, 1–16.

333. Environment Canada Daily Data Report for February 2019 - Climate - Environment and Climate Change Canada.

Appendix A. Supplementary figures

>TpTS1 (OM324392)

CTGAAAGCTCCTCAATATTCTGCACTCCACTTTTCTGTATATAAAAAGCGCAGCAAACAAGCATTCCTGGC
ATCCTCCTTTTCTACAACTTAGAAACCATTTCTATCTGAGAAAGGGGAGGAAACAAATAAGTTGTTGTGCA
ATGGCTCTTTCCTCTGCTTTTTCATACAATTTCTGCCTCAAATCTCAGCCTTCCCAATCCCAATCCCAATC
CCAAACCCAGTCCCACACTTTTGCTGTTAACAAAGGTAAAGCAGGGCTGCCATTAAAGCCCAGAACAAAGG
TGTTCATGAATAACCCATTAATAGTAAATCAGGCCCAAATCACCAAAACTGAAATCCCAGAGCGGCGCATT
GGAAACCATCATCCCAATTTGTGGGGTGATGATTTTATACAGTCTCTCCCAGACCATCCTTATCAGGGTGA
TCACTATGCTGAACGTTGTGGGAAACTTATTAGCGAGGTAAAACACATGTTTATTGTTGCTGAAAAAGATG
TTTCACAACTTCTTTTCCTGGTGGATAACATCGAACGCCTTGGAATAGATCGACACTTTCAAAAGGAAATA
AAACAATCCCTTGATTTGGTTTACAGGAATTGGGGTGAATGCCAAAGAGATCTCAACACCACGGCCTTGAG
CTTTCGGATTCTCAGGCTCCATAGATACTCTGTTTCTCCAGGAGTATTAGGACCCTTTAGAACAACAAGTG
GGCAGTTCTTGTGCTCCACCGCTCAGTCAGAGGAGGAGAAAATTAAAAGTGTTCTGAATTTGTACAGAGCT
TCATATATTGCATTTCCAGGGGAGAAAATTTTGGATGAGGCCAAACAATTCTCTACTTCATATTTAACTGA
AGCTCTGCACAAGACAGGGATCAACTCCAGCCTTTTGTATGAGATATCAGTGAATCTGAGGTATGAATGGT
ACACACTTCTGCCCAGGTTGGAAGCAAGGAAATATATAGAAATCTATGGGGAAAACACTTCATGGGCCAGG
ATGGCGGATAATGAGAAGCTTTTATCCCTAGCGAAAATGGATTTCAACATGATTCAGTCATTGCATCAGCA
AGAGCTTAAAGCTTTCTCCAGGTGGTGGAGGGAGTCTGGTTTGCCTAATGTAGATTTTGCTCGTCATCGAC
ACGTAGAATTTTTCTTTATGGCATGTGCAATTTGTGAGGATGAGAAGTATTCTGCATTTAGATCGAGCCTT
GCGAAGTCTTGTGTACTTGCAACATATTTAGATGATACATATGACACATATGGAACAGTTGATGAACTAGA
ACTCCTCACACAAGCAATTAAGAAGTGGGATCCGGCGTCTATAAACCAGCTCCCTGAATATATGAAAATTA
TATACATGGTTTTGTATAATGGAATCAATGAAATGGATCGAGAAGCTCAAAAGTTTCAAGGCCGAGAGACT
CTCTCCCATGCTAAAGATGCTTGGGAAACTTATATAGATGCTATGCTGCAAGAAGCTAAGTGGAATGCTAC
AAAACACATACCAACTTTGGCAGAACACATAGAAAATGGAACGATTAGTTCAGGAAAACGTCCAACCACAT
TGCAAGCTATATTAACAGCAGATGGGATTGTTCCAGAAGATATCTTTCATAAAATTGATTATCCATCAAGG
TTTGATAATCTTGGAGGCTTCAGCCTTCGAGTAAGAGGTGACATCAAATCCTTTAAGGGTGAGGTGGATCG
AGGAGAAACAAATTCATCCATAGTGTCCTACATGAAAGATAATCCAGGATCCACTGAAGAAGATGCCTTAA
ATTATCTTGTGAATTTATTGGATGAAAGACACAAAGAATTGAATTGGGAGTTTCTAAAAAATGATGGTGTT
CCAACTTGTAGTAAGGATTATTCTTATGATGTTGCAAGGGGTTACCTACATTTCTACAATGAGAGAGATGG
GTTTAGTTTTGCCTATCAAGAAATACGAGACCATGTAAATCAAATCCTCATTGAGCCTGTATCTATGTAAC
TGCTCATTTTAGTGAACCAATGTCTACAATCACATAAGTCTCTCTTTACTTGACTGAGAGACATTTTTTTA
TTATCAATAATAAATGAATAGGCGTTTCTTTCTAAAAAAAAAAAAAAAA

>TpTS1

MALSSAFSYNFCLKSQPSQSQSQSQTQSHTFAVNKGKAGLPLKPRTKVFMNNPLIVNQAQ
ITKTEIPERRIGNHHPNLWGDDFIQSLPDHPYQGDHYAERCGKLISEVKHMFIVAEKDVS
QLLFLVDNIERLGIDRHFQKEIKQSLDLVYRNWGECQRDLNTTALSFRILRLHRYSVSPG
VLGPFRTTSGQFLCSTAQSEEEKIKSVLNLYRASYIAFPGEKILDEAKQFSTSYLTEALH
KTGINSSLLYEISVNLRYEWYTLLPRLEARKYIEIYGENTSWARMADNEKLLSLAKMDFN
MIQSLHQQELKAFSRWWRESGLPNVDFARHRHVEFFFMACAICEDEKYSAFRSSLAKSCV
LATYLDDTYDTYGTVDELELLTQAIKKWDPASINQLPEYMKIIYMVLYNGINEMDREAQK
FQGRETLSHAKDAWETYIDAMLQEAKWNATKHIPTLAEHIENGTISSGKRPTTLQAILTA
DGIVPEDIFHKIDYPSRFDNLGGFSLRVRGDIKSFKGEVDRGETNSSIVSYMKDNPGSTE
EDALNYLVNLLDERHKELNWEFLKNDGVPTCSKDYSYDVARGYLHFYNERDGFSFAYQEI
RDHVNQILIEPVSM

>TpCS1 (OM324393)

CGCCAGAACAAAGAGAAAGTTCCTTCTTGCTCAGAAGAATAGTATTAAAGAGCAAAAATGGCTGTGAGTTG
CATTACGCCTTTGGCTTCCACCATGGTTGGACAGAAGCTACCACTTGGTAGAGTACCTTCTATTCAGCGTA
AGCCTTTCAGCAGAGCATACCCAGCACGTCGCAGTGTTGCTCAGCTTCCAATCAATATGGCGTTGATAACC
GATGATAAAGGCATTACACGGCGCATTGGCAATCATCACCCGAACTTGTGGGATGATGATTTCCTGCAATC
CCTCTCAAAATCTTCTGGCGCTCCCTCAGATGGTGAACGTGCTGAAAAGCTCATAAATGAGATAAAGGGCA
TATTCAATGCCCTTCATCTAAACTCTTCATCTACAGACGATCTCTATCGACGTCTTTCAATGGTGGATGAT
GTGGAACGCCTTGGAATTGACCGCCATTTTCAAAATGAAATAAGAGAACATCTTGACTATGTTTACAGATA
TTGGAATGATACCAATGGCATTGGGTGTGGTAGAGAGAGTCCTTGCGTGGATCTCAACGCCACAGCCTTAG
GCTTTAGAACTCTGCGCCTCCACAGATATGGCGTCTCTTCAGGTATGCTGCAACAGCTCATAAGTAAAGAT
GTACATTTTTTGTATAGTCAATCGAATGAGAAGGAGCTAAGAAGCATTCTAAATTTATTTCGAGCTTCACT
CATTGCATTTCCTAAGGAGCAAGCTTTGGATGATGCCAAAGCCTTTTCTACTTCATATTTAAAAAATGCTC
TAGCCAATATTAATAATACCAATCTTTCAAAAGAGATACAATTTAATCTAGAGTATGGATGGCACACTAAT
GTGCCTAGGTTGGAAGCAAGGACATATATTGACATATATGGAGATGTCAAAACTACAAACAATATTTTCAA
CTCGAAACTTTTAGAACTAGCCAAGATAGACTTCAATAAGATTCAATCAATTCAACAACAAGAGCTTCAAA
TTCTCTCAAGATGGTGGACAGAGTCTGGTCTAGCTAAACTAGAGTTTGCTCGCCATCGCCATGTGGAATTT
TACTTTTGGGCAGCAGGAGTGTGTATTGAACCAAAATATTCTACATTCAGAATTGGGTTTGCAAAGATGGC
TACTTTTGTCACTTGTTTGGATGATATTTATGACACTTATGGAACTATAGGTGAACTTGAAATCTTCACAA
AAGCCATTAAGAGGTGGGATTTACGATCCATGGAAGGACTTCCAGAATTTATGAAGATAACATTCAAGGCC
TTTGACGAGGGTGTGAGGGACATGGCTCGAGAGGCTGAAAAGACTCAAGGCAGAGACACACTTGACTATGC
GCGAAAAGCTTGGGAAGTTTACATAGATGCTTACTTACAAGAAGCAAGGTGGCTTGCTATGGGCTATATAC
CATCTTTGGAGGAATATCTAGAGAATGGAAAAGTAAGCTCTGGGGCTCGTGTTGTGACCTTACAACCCGTA
CTCTCATTAGATGCTCCCATTTCTGAGGATGTTATTAAAGAACTTGATTATCCATCAAGGTTCAATGAAAT
TTTATGCTTAACCCTTAGATTAAGAGGTGACACAAAGACTTTCAAGGCTGAGGCAGATCGTGGAGAAGAAG
TATCATGTATAACATGTTACATAAAAGACCATCCTGGATCTACAGAGCGAGATGCATTGGAGTATCTCAAT
AATTTAATTGATGAAAGACTTAAAGAATTGAATTGGGAATATTTAAAACCTGATAATGTTCCAATAATCAG
CAAAGATCATGCATATGGCTTATCAAAAGGTCTCCAGCTTCTCTATAAAGAAAGAGATGGATTTAGTGTGT
CAGGCATTGAGACAAAGAATATTATCACCAAAACCATGATTGAACCAGTTCCTATGTGACTATTTAATTAA
GTTATATATACTTGTTATCAAGAATAGATTATTGTTTGATACATTTAGTCTATCATAATGACTATGGTTTA
TTAATCATCTATGTAATCAATGTATATACACACTTGTGTGTGTTTATAAATAATGATTTTCTTTCAAAA

>TpCS1

MAVSCITPLASTMVGQKLPLGRVPSIQRKPFSRAYPARRSVAQLPINMALITDDKGITRR
IGNHHPNLWDDDFLQSLSKSSGAPSDGERAEKLINEIKGIFNALHLNSSSTDDLYRRLSM
VDDVERLGIDRHFQNEIREHLDYVYRYWNDTNGIGCGRESPCVDLNATALGFRTLRLHRY
GVSSGMLQQLISKDVHFLYSQSNEKELRSILNLFRASLIAFPKEQALDDAKAFSTSYLKN
ALANINNTNLSKEIQFNLEYGWHTNVPRLEARTYIDIYGDVKTTNNIFNSKLLELAKIDF
NKIQSIQQQELQILSRWWTESGLAKLEFARHRHVEFYFWAAGVCIEPKYSTFRIGFAKMA
TFVTCLDDIYDTYGTIGELEIFTKAIKRWDLRSMEGLPEFMKITFKAFDEGVRDMAREAE
KTQGRDTLDYARKAWEVYIDAYLQEARWLAMGYIPSLEEYLENGKVSSGARVVTLQPVLS
LDAPISEDVIKELDYPSRFNEILCLTLRLRGDTKTFKAEADRGEEVSCITCYIKDHPGST
ERDALEYLNNLIDERLKELNWEYLKPDNVPIISKDHAYGLSKGLQLLYKERDGFSVSGIE
TKNIITKTMIEPVPM

>TpES1

AGGAGATGTAATCGATCGGCTTAAGATGGTTGATCRCAAGCATTCAACTTGTAAATTTGAGTTGTAGTCCA
CATTTCTTCCGTGATGTCTAATTTGAAGGGAGACCAGACTTCCATGGAAATGCCAGAAAAATCCTTAAATC
TGTGGGACGACGAGTTTGTCCAATCTATGGAAACACCATATAATTTATGGGCATCCGAATATCCTGAACGT

GTTGAAACACTGGTTAAACAAGTCAAAATCTTGTTTAATGAAATGCAGAGTGGAGGAGGAGATGTAATCGA
TCGGCTTAAGATGGTTGACGCATTTCAATGCCTCGGAATAGATCGATATTTTGAGGATGAAATAAAAGCGG
CTCTTGATTACGTTTACCGCTGTTGGGATGGAAGTGTGGGGATAGGATTAGGCTGCGAAAGCGCTACAAAG
GATTTGAATGCCACGGCTTTGGGACTCAGAGTGCTTCGACTCTATCGTTATGATGTGTCTGCAGATGTGTT
GAAGAACTTCAAGGACAATGAAGGGCAGTTCATTTCCCGTGGTGATGAGGAACACATGATGACAAGTATGA
TCAACCTCTTAAGAGCTTCAAGTATGTCATTTCTTGGAGAGGTGGTTATGGAAGACGCTAAAGTCTTCAGC
TCTGCTTATCTTAGACAATTATTAGAGAAATCTCAAGATATAAAGCACAACAGTTTTCTGAAAGAGGCTGA
GTATGCCCTCCAATATGAATGGCCTTCTACTTTCTCTCGATGGGAGGCCCGAAGTTATATAGAAATATATG
AATTGGATGTCTCCAGGTTGAAGGACAAAATCGTTTTAGATTTGGCCAAATTGGATTTCAATATATTGCAA
TATATATACAAAATGGAAACGAAGATGCTCTCCAGTTGGTGGCAAAATTCTGGTGTCTCAAAACTGATTGC
AGTTAGGGAGCGCTCAATTGAATATTATCTGTTGGCAGTTAGTGCTGTTGATAATGCAGAGTTTGGTAGCA
GCAGAATAGCTTTGGCCAAAGCAGCAACTCTCGTCTCTCTATTAGATGATCTTTTTGATGATTACTTGACA
CTTGAGCAAGTTGAACTTGTTACCAAGGCTATTGTTGAAGGTTGGAGCATTTCCATTATACAAAATATTCC
AGACAATTATAAGAAAATTGTGGAATTTATTTTCAAAACATTACATGAATTGACAAGTGAGGCAACTCAAA
TTCAAGGACGAGAGATGATGCACTTCATTACAAAAGCGTGGGCCGATTATGCTAAAGCTAGTTTAAAGACA
GCACAATGGAAAGAAGGTCAATATGTTCCAACCTATAATGAGTATATAAAGGTTGCTACCACAACTGGGGC
AAGTGGACTATTATCATTGCATCCCATCCTCCTGGCAGTTCCTAATTTGGAAGATGATGCTATTGAGAAAA
TATTTCTCAATAAGTCAAGATTCTATGAGCTTGTATGGTTGACGGGACGATTAGTTGATGATGTTCATGAC
TTCCAGGATGACAAGCTCCATGGACAAACAGCGTCAGCAATTTCTTCTTATATGAAGGACCATCCTGAATG
TTCAGAAGAAGAGGCATTACTTCATATCAATAGTCTTTTTGACCAATTGCTTAAAGAATTAACACTGGAAT
ATTTGAATCCCATCAAAGTGTTACCCAAATGGGACAATTTATACTTCAATCTTGGGAGAGGAGTTCAAAGC
TTTTATGTATTTGGAGATGGATTTTCATATCATGATAAGGGAGTCAAGCAACGAGTATTCAAAGTTCTTTT
TGATCTAGTTAAAGTTTAAAAAAATAGTGTTGCTATAATTGTAAGTTATGACTCTTGGTTGCCATACTTGG
CTTCCAATATGAGGTAGTTTCTAAATAACAGTGTATTTCAAAGGGTTGACTCTTAGTCCTCCATAGATTTA
TCCTACTGTGTTTTTCTCATAGTTGACATTTAATTTTTTTAAAATACTATATTTCTACTTTCTGGTTAAAA
AAAAA

>TpES1

MSNLKGDQTSMEMPEKSLNLWDDEFVQSMETPYNLWASEYPERVETLVKQVKILFNEMQS
GGGDVIDRLKMVDAFQCLGIDRYFEDEIKAALDYVYRCWDGSVGIGLGCESATKDLNATA
LGLRVLRLYRYDVSADVLKNFKDNEGQFISRGDEEHMMTSMINLLRASSMSFLGEVVMED
AKVFSSAYLRQLLEKSQDIKHNSFLKEAEYALQYEWPSTFSRWEARSYIEIYELDVSRLK
DKIVLDLAKLDFNILQYIYKMETKMLSSWWQNSGVSKLIAVRERSIEYYLLAVSAVDNAE
FGSSRIALAKAATLVSLLDDLFDDYLTLEQVELVTKAIVEGWSISIIQNIPDNYKKIVEF
IFKTLHELTSEATQIQGREMMHFITKAWADYAKASLKTAQWKEGQYVPTYNEYIKVATTT
GASGLLSLHPILLAVPNLEDDAIEKIFLNKSRFYELVWLTGRLVDDVHDFQDDKLHGQTA
SAISSYMKDHPECSEEEALLHINSLFDQLLKELTLEYLNPIKVLPKWDNLYFNLGRGVQS
FYVFGDGFSYHDKGVKQRVFKVLFDLVKV

Figure A.1. Nucleotide and predicted amino acid sequences of two *Thuja plicata* monoterpene synthases and sesquiterpene synthase described in Chapter2.

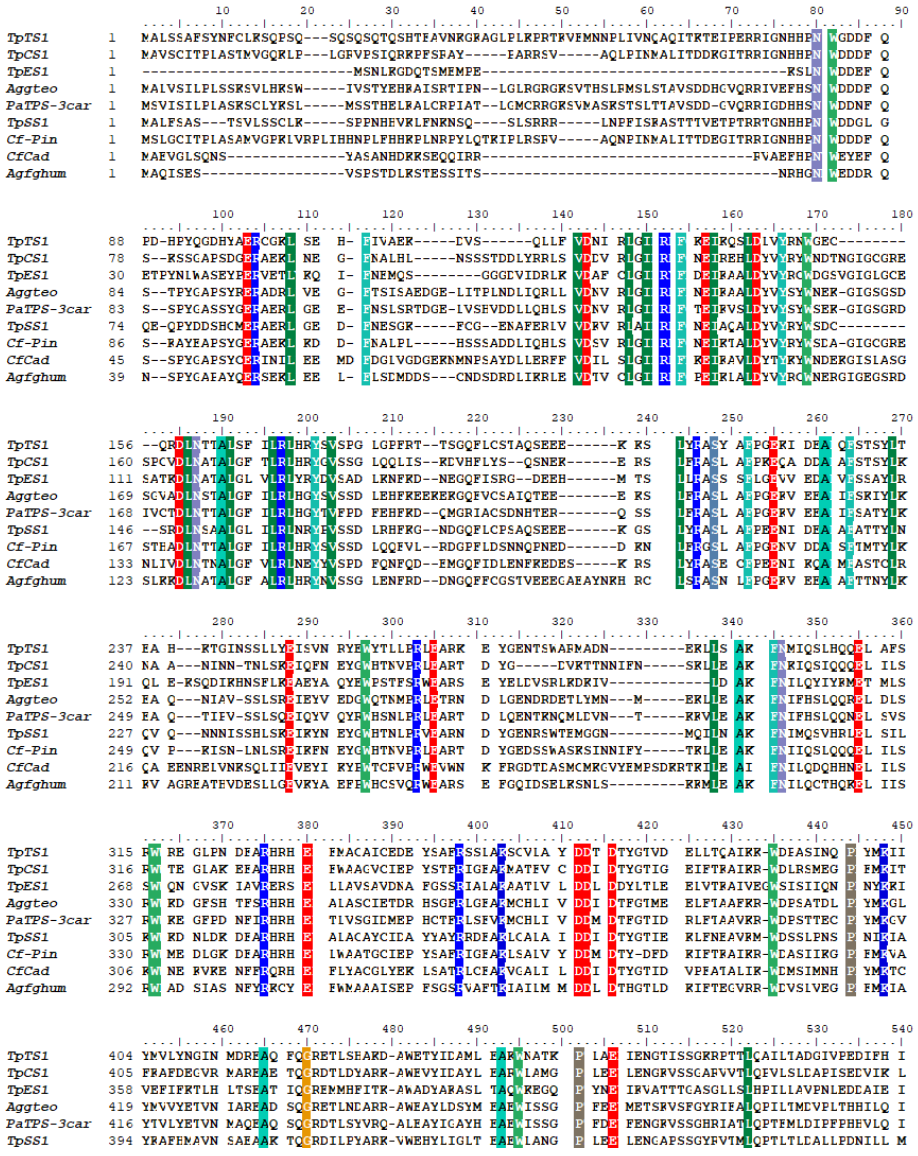
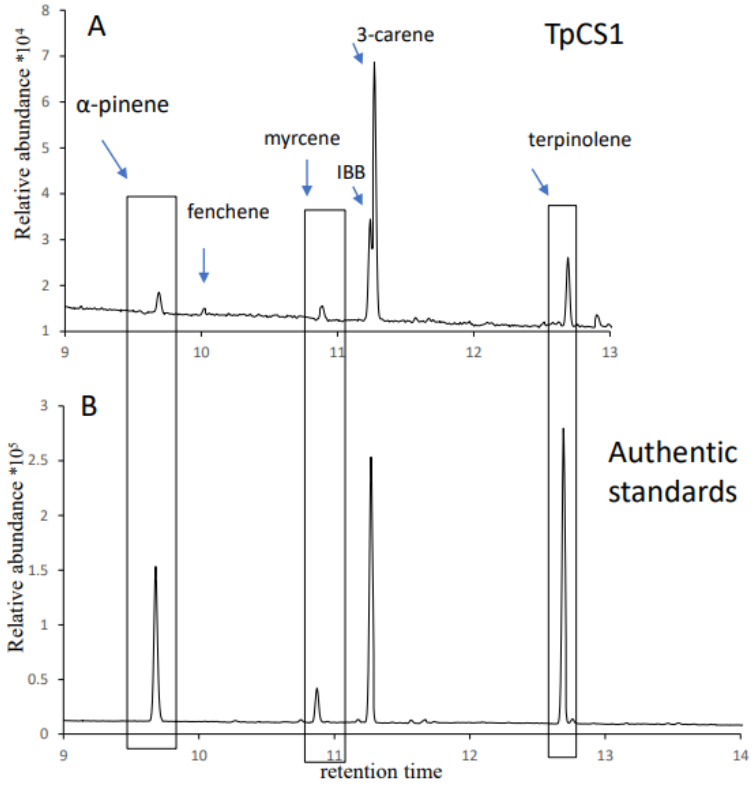


Figure A.2. Multiple protein sequence alignment of *Thuja plicata* TpTS1, TpCS1 and TpES1 with *Pseudotsuga menziesii* terpinolene synthase (PmeTPS2, AAX07264.1), *Picea abies* (+)-3-carene synthase (PaJF67, AA073863.1),*Thuja plicata* sabinene synthase (TpSS1, KC767281.1), *Chamaecyparis formosensis* alpha pinene synthase (Cf-Pin, EU099434.1).


A
TpCS1
3-carene
α-pinene
myrcene
IBB
terpinolene
fenchene
Relative abundance *10⁴
B
Authentic standards
Relative abundance *10⁵
retention time

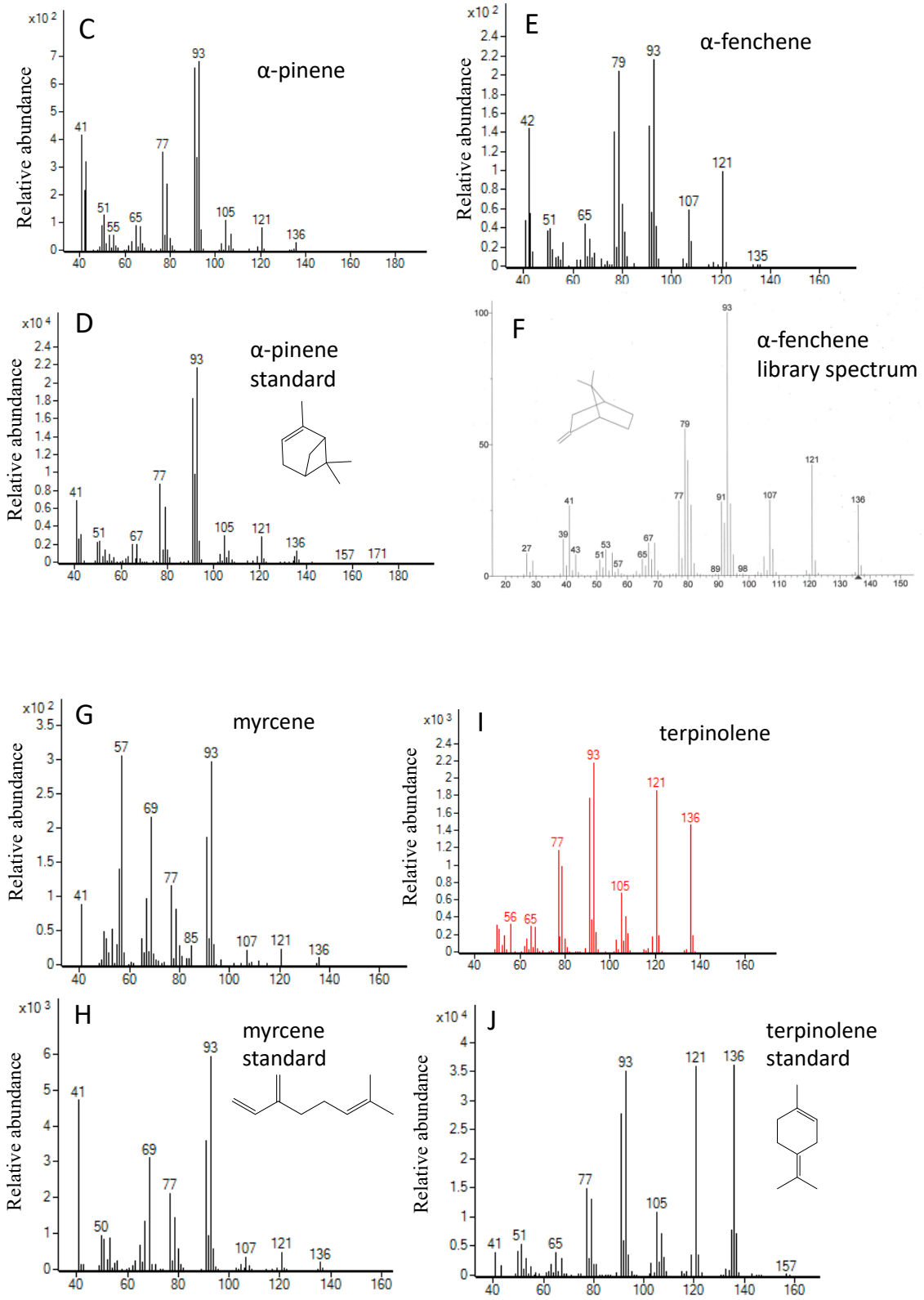


Figure A.3. (A, B) = total ion chromatogram (TIC) of terpenes produced by TpCS1 and authentic standards (C, D) = mass spectrum of α-pinene and standard α-pinene, (E, F) = mass spectrum of

α-fenchene and library spectrum. (G, H) = mass spectrum of myrcene and standard myrcene. (I, J) = mass spectrum of terpinolene and standard terpinolene.

>TpdiTPS1 (MT468207)

CTTCTAATGCCTACTGAAACAAATTTGATCTCTATTTTATCAAAATGGCTCAGAAGATGTTTTCTCCGTCC
ATTTCAGTTTCCAAGTCTCGAGGATGGATTTCAACCAAAACCTCTGGATTTTCACCCATTGGAAAGTACTC
GCGCTCAAAGAGTCTGGCTTGCCACAACATGCCATCTGCTGTTGTTGTTGGAGAGGATGCTAAGACTCTTC
AGGCTTTGAGAATAGCGCATAAGGAGAGTAAAATCAATCCTAATCACGCGAAACCAGATTATGTTCACTCT
GTTTCAACATTTGAGGAGGCACCCTTTGATGAGATGGATAAGAGAATAGAGGAATTGGTTACAGAGATCAA
AGGGCTTTTTAATTCGATGGAAGATGGGGAAATAAGTCCCTCTGCTTACGATACTGCGTGGGTAGCGAGAG
TTCATGCCATTGATGGCTCTGTTAAACCCCAATTTCCTCAAATGGTGGACTGGATTCTTCAAAATCAGTTA
CCGGACGGTTCTTGGGGAGAAAAGAGTCGCTTTTTAGCGTGTGATAGACTGCTCAATACTCTTTCCTGCTT
AGTTACTCTCTCCATCTGGGGCGTTGGAAACAATCAAGTGAACAGGGGTCTTAATTTCTTGAGGAGAAATA
CAGAAGGAATGATTAAAGAAGCACTCGGTCATCATCAACCCAAGGGATTCGACATGGTCTTTCCTGTACTG
CTGAATGAAGCCAAACTTTTGGGTTTGGATCTTCCTTATGGGCTATATATCGTTAAACAAATACGCGAAAA
GCCAGACTTGGAATTGAAAAAAGTATTTGTTGAAGAGCTACACGGTCATCCCTCAACAATGTTACAGTGTC
TGGAAGGCGTACAAGAAATAATTGATTGGAAAAGGGTCCTGAAACTGCAATCCAAGGATGGATCTTTCTCT
GGCTCGCCAGCATCTACAGCTTGCGTATTTATGCACACAGGAGACATTAAATCTCTCCAATTTTTGACGAG
CCTTGTTAAAAAGTTTGGAGACCATGTGCCGAGCATGTATCCGGTGGATATAGCGGAGCGTTTAAGAGCCG
TGGATTGCGTTGAGCGTCTTGGACTCGAACGCCATTTTCAAACGGAAATTAAACAAGCAATGGATTATGTG
TTCCAGTACTGGAGTGAAAGAGGCATTGGATTTGGAAGGGAAAGCCTGGTTCCTGATATTGATATTACAGC
CACTGCCTTCAGGCTTCTCAGAACCTTCGGCTACTCTGTATCTTCCGAGGTTCTGCAAAACATTAAAGCCG
AGGCTGAAGAACTTCTTAAGCTGTCTGGTAACGAAAACAGCGCAGGAATAATCGGAATATTGAGCCTGTAT
AGATCTTCACAGCTTAACTTTCCGGGCGAAATTGTAATGAAGGAAATAGATTGCTTTGCAAAAGATTACCT
GGCTGAATTCCTGCAAACCAAGAACTTTTCTCAGGTGAAGGTTGTCAAGGAAAACCTTCCCGAAGAGGTTG
AATATGCTTTGTCTGCTCAATGGAATAGAAATATGCCAAGGCTGATGTCTAGAAATCAGATAGAATTGTTT
AATCCCAATGACCTATGGCTGGGGAAGACATTATATCACCTGCCAAATGCTAGCAATGACAAGTATTTAGA
ATTGGCCAAACTCGACTTCAACCGCATTCAGGCTACACACAGATTTGAAATACAACGAATACAGAGATGGT
ACAAGGATTGCAATTTCCCGCGGTTGGATTTTACTCGTCACAGAGAAGTGGCAGTTTACTGGACCTCTTCA
GCGGTGATGTTTGAACCACAATACACCGATTGCAGACTCGATTACGCAAAAGCCGGACTCCTGGCTACTAT
CACAGACGACCTCTATGAATCCTATGCAACTCTCAACCAACTCAAGCTTTTCAATGAAGCTTTCGAAAGAT
GGGATCCATTAATGAGCGAGCAACTGCCAGAGGACATGAAAATAGTTTTCATGGGAATCTATAACACTTTG
ACCGATATTTCTGAGCGAGCACTGAAGGTTCAAGGTCGTGATGTGCTTCCTTACCTGCGCCAACAGTGGTT
AAATCTGTTGTTCAGTTTCACAAAGGAAAGAGAGTGGATGGAAAGAAGTTATTCGCCATCATTGGATGAAT
ACTGGGCGAATGCAGAAGTGTCGATAGCACTGGAAACAACTATCCTGTCGCCAATATTTTCTACGGGAAAT
TTTCTTCCTGATCGCATTCTGGAGAAGATTAACTTTCTGGATCTGGTCAGTCAAACAGGGCGTCTTATGAA
CGACGTTAGAACTTTCCAGAAGGAAAGAGATCGCGGGGAATTGGCTTCATTCGTAGAATGCTACAAAAATG
AGCTCCATGGATGCACTGAAGAAGAAGCCCTGAATTACATGGAGAGAATGAACGAAGAGGCCCTTATCAAT
TTGAACTACCATTTTCTAATGCGATCAGACATACCTAAGTGCTACAGAACACTTCTTTTCAACACGGCTAG
GATAATGCAAATGATTTACAGGAAGGAAGACGGTTTTCGAAATGCCGCGGAGTACCTCGAAGATTCAATTA
AAAAATCTCTGTACGAACCAGTGCTGTAACAGATTACTATTCTGTAGTGTTGAAATTCGAACATCAGTCAA
ATAAAAAAAAATTATTCCAAATGTTAATAAAATATTTTTTATTACTAAAAAAGGCATCATTACAATCATGGT
AATATTATTAGTTAATTTTCTGCAAAAAAAAAA

>TpdiTPS1

MAQKMFSPSISVSKSRGWISTKTSGFSPIGKYSRSKSLACHNMPSAVVVGEDAKTLQALRIAHKESKINPN
HAKPDYVHSVSTFEEAPFDEMDKRIEELVTEIKGLFNSMEDGEISPSAYDTAWVARVHAIDGSVKPQFPQM
VDWILQNQLPDGSWGEKSRFLACDRLLNTLSCLVTLSIWGVGNNQVNRGLNFLRRNTEGMIKEALGHHQPK
GFDMVFPVLLNEAKLLGLDLPYGLYIVKQIREKPDLELKKVFVEELHGHPSTMLQCLEGVQEIIDWKRVLK

LQSKDGSFSGSPASTACVFMHTGDIKSLQFLTSLVKKFGDHVPSMYPVDIAERLRAVDCVERLGLERHFQT
EIKQAMDYVFQYWSERGIGFGRESLVPDIDITATAFRLLRTFGYSVSSEVLQNIKAEAEELLKLSGNENSA
GIIGILSLYRSSQLNFPGEIVMKEIDCFAKDYLAEFLQTKNFSQVKVVKENLPEEVEYALSAQWNRNMPRL
MSRNQIELFNPNDLWLGKTLYHLPNASNDKYLELAKLDFNRIQATHRFEIQRIQRWYKDCNFPRLDFTRHR
EVAVYWTSSAVMFEPQYTDCRLDYAKAGLLATITDDLYESYATLNQLKLFNEAFERWDPLMSEQLPEDMKI
VFMGIYNTLTDISERALKVQGRDVLPYLRQQWLNLLFSFTKEREWMERSYSPSLDEYWANAEVSIALETTI
LSPIFSTGNFLPDRILEKINFLDLVSQTGRLMNDVRTFQKERDRGELASFVECYKNELHGCTEEEALNYME
RMNEEALINLNYHFLMRSDIPKCYRTLLFNTARIMQMIYRKEDGFRNAAEYLEDSIKKSLYEPVL

>TpdiTPS2 (MT468208)

GGAATCTAATTGAGATTTGGCAGGGTATGCCCTGTGCAGCTCCGATAAAACCCTCTCTCTCTCTACTCCCC
ATGGGATTCAAAGTCTTCTTCACAAGGCGACATCTAATTTGTTAAAAATGTCCCAGAGTTTGTGTCCGCGC
CTCAGTTTGTTTTCGAAGCCTACTACCAAATCCACACAACGTCTTTCCAATACTTCTCTGCCTTTTACAAA
CTACGCTCGCATTAAGAGCATAGATTGCTACAACATGACATCTGCTCCTGCTCTTGGCGACAATGCGAAAA
CACTTCACGCTGCAGCGATTGCGCATCCGGAGCCCAAAATCTATCCTAATGCTGGGAAGCCAGATTATGTT
CATTCTAATTCAACATTTGAGGAGGCACCGTTGGAGGAGATGGATAAGCGAATAGAGGCATTGGTTGCGGA
GATCAAAGAGCTGTTTTATTCAATGGAAGATGGGGAAATAAGTCCTTCCGCATACGATACTGCGTGGGTAG
CGAGAGTGCCTGCCATTGATGCCTCCGCTCAACCCCAATTTCCCCAATTGCTGGACTGGATTCTTCAGAAT
CAGTTAGCGGACGGTTCCTGGGGCCAGCAGAGTCGCTTTTTAGCGTCTGATAGGTTCCTCAATACTCTTGC
CTGCCTCCTTACTCTCACCTTCTGGGGCGTTGGAAACAATCAAGTGCAGAGAGGTCTTCATTTCTTAAGAG
GAAATATGGAAGCAATGGTTAAAGAAGCTGTAGCATTCGGTCATCAAGGATTCGAGATGGTTTTGCCTGCA
CTGTTGAATGAAGCCAAACATTTGGGCTTGGATCTTCCTTACGAGCTACCTATCATCCAGCAAATAAACAA
AAAGAGGGACTCCGAATTGAAAAAGGTATCTGTTGAGGAGCTACACACGCATCCGACAGCAATGTTGCAGT
GTCTGGAAAGCATACAAGAAGTAGTCGATTGGAAAGACATCCTGAAATTGCAATCGAAGGATGGGTCTTTC
TCAGGCTCGCCAGCATCTACAGCTTGTGTATTTATGCACACCGGAGATAAGAAATGCCTACGATTCTTGGC
GGGTCTTGTTACAAAGTTTGAAGACTATGTCCCCTGCATGTATCCAGTGGACATAGCAGAGCGTTTGAGGG
CCGTGGATAGTGTTGAACGTCTGGGGCTTGAACGCCATTTCCAAACGGAGATCAAACAAGCCTTGGACTAT
GTGTTCCAGTACTGGGGTGAAAGAGGAGTTGGATTTGGAAGGGACAGCCTGGTTCCGGATATTGATGTCAC
AGCCACGGGCTTCAGGCTTCTCAGGATGTTCGGCTACACTGTGTCTCCAGACTTTCTGCAAAACATCAAAG
ACGAAGCTGAAGAACTCTGTAAGCTGTCTGATGGTGAAAACAGGGGAAGAGTAATCGACATGCTGAGCCTG
TATAGATGCTCACAGATTAACTTTCCGGGAGAAAATGTAATGAGAGAAATAGGTGCATTTGCCAAAGATTA
CTTGGCCGAATCCCTGCAAAGCAACAACTTTTCTCAGGCGACGGCTGTCAAGGATAACCTTCGCCAAGAGG
TCGAATATGCTTTGTTTGCCAGATGGAATAGAAATATGCCGAGACTGGTGATTATAAATAACATAGAAGTG
TTTAATCCCGATGACTTATGGCTGGGGAAGACATTATATCAAATGCCAAATGCGAGCAACGGCAAGTATTT
AGAATTGGCCAAACTCGAGTTCAACCGCACTCAGGCTATACACAGATCCGAAATACAACATATTAAGAGAT
GGTACAAGGCTTGCAATTTCCCCCAGCTGGAATTTACTCGTCACAGAGAAGTGGCAATCTACTGGACTGCG
GCAGCGGTAATGCCCGATCCCCAATACACCGACTGTAGACTCGCTTATGCAAAAGCAGGAATCATGGCTGT
TATCACAGACGACTTGTATGACACCTGTGCAACTCTGGAGCAGGCCAAGCTCTTCAACGAAGCTTTTGAAA
GGTCGCCGCGTTTAATCCCCATCTCAGTTAATGCGAATAAGGAATTGGAATCGGCGATATTAACATTGATT
GTGTTTGATGGGTGCAGATGGGATACAGAGCAAATCGAGCATCTACCAGAGGAGATGAGAATAGTATTTAT
GGGGCTGTACAACACTTTGAGGGAGATATCTGAAGGAGCGCGGGAGGTCCAAGGGCGTGATGTGCTTCCTT
ACCTGCGGCAAAAGTGGTTGGATCTGTTCAGGAGATACACGAAAGAAACAGAGTGGATGGAAAGGAGGCAT
TCGCCGTCATTAGAAGAATACTGGGAGAACGCGGTGGAGTCGATAGCACTGGGAGTCACTACCCTCACCCC
AATATTCTCTACCCAAGATCTTCTTCCTGATCATCTCCTCCAGAAATTTGACTTCCGCGCAGACTTTCTGA
ATCTCGTCAGTCTCACGGGGCGCCTCATTAACGACGTCAGAACTTTCCAGGAGGAAAGAGATCGGGGGGAA
TTGGCTTCATGCGTGCAATGCTACAGAAATGATAATCCGGGATGCACGGAAGAAGAAGCCCTGAATTATCT
GTACGGGGTGAATGAGGACGCCCTGACTAAATTGAATTATCAGTTTTTGATGCGTGAAGACATTCCCAAGA
GCTTCAGAACTGTTCTTTTCAACACGGCCAGGGTAATGCAATTGTTTTACAGGAACATTGACGGCTTTCTA
AATGCCGCCGAAGAGATGAAAGTCTTCATTAAAAAGACTCTCTATGAACCCCTGCTCTAATAAATTGCTCT
ATTGGATGAACTGTGTAATAAATGAAATTCCATTTTTTTTTTAGTTTTTGGTGAATATATTTGAAGGAGAA
ACTTATTAATGTACACTTCAATTTATGAAAGGATAACGAATCAAAATTTACTATCATT

>TpdiTPS2

MSQSLCPRLSLFSKPTTKSTQRLSNTSLPFTNYARIKSIDCYNMTSAPALGDNAKTLHAAAIAHPEPKIYP
NAGKPDYVHSNSTFEEAPLEEMDKRIEALVAEIKELFYSMEDGEISPSAYDTAWVARVPAIDASAQPQFPQ
LLDWILQNQLADGSWGQQSRFLASDRFLNTLACLLTLTFWGVGNNQVQRGLHFLRGNMEAMVKEAVAFGHQ
GFEMVLPALLNEAKHLGLDLPYELPIIQQINKKRDSELKKVSVEELHTHPTAMLQCLESIQEVVDWKDILK
LQSKDGSFSGSPASTACVFMHTGDKKCLRFLAGLVTKFEDYVPCMYPVDIAERLRAVDSVERLGLERHFQT
EIKQALDYVFQYWGERGVGFGRDSLVPDIDVTATGFRLLRMFGYTVSPDFLQNIKDEAEELCKLSDGENRG
RVIDMLSLYRCSQINFPGENVMREIGAFAKDYLAESLQSNNFSQATAVKDNLRQEVEYALFARWNRNMPRL
VIINNIEVFNPDDLWLGKTLYQMPNASNGKYLELAKLEFNRTQAIHRSEIQHIKRWYKACNFPQLEFTRHR
EVAIYWTAAAVMPDPQYTDCRLAYAKAGIMAVITDDLYDTCATLEQAKLFNEAFERSPRLIPISVNANKEL
ESAILTLIVFDGCRWDTEQIEHLPEEMRIVFMGLYNTLREISEGAREVQGRDVLPYLRQKWLDLFRRYTKE
TEWMERRHSPSLEEYWENAVESIALGVTTLTPIFSTQDLLPDHLLQKFDFRADFLNLVSLTGRLINDVRTF
QEERDRGELASCVQCYRNDNPGCTEEEALNYLYGVNEDALTKLNYQFLMREDIPKSFRTVLFNTARVMQLF
YRNIDGFLNAAEEMKVFIKKTLYEPLL

>TpdiTPS3(MT468209)

CTCTTAACACCTATTCATTGAGGTCGCTGACAAATTTGAAAGGCAGGTGCAATTGGGACTTATCAGCTGAA
CAGTGCAAAAGCTTAAAAACATACCAGTATTTTTTCTTCTTCTTCTTTTCTAACTTAGAGAAAATGGCTCA
ATCTTTGATTTCTTCGGCCAATTCTTGTATTGCAAAGACTCGACTTTCAGATGTTAATGGCAGAACAGAGC
TCAACAGATCTCGACCATTATCATTTTCATTTTCCTCCTCTAAGCTGGTGAAGAGGCAATGGAATAAGCAG
TCAGCTGGTTTGGTTTTGGCTTGTCTTGGAGATTCCAGGCGTTCTTCTTCCCCTGCTGCTACAGGGGCTGC
CACTGCTTCTACATCGGTGAAACGAGAGTACCCTCCAGCAGTTTGGAATGATGACGTGATCAATTCTCTGA
TATCAACTTATAAATCTGCAGATGTAGCAGAGCAGGAGAAAAGGTCTGAGACATTGATAGCGGAGATAAAG
GGCATGTTTAAATCAATGGGGGATGGTGAAACGAATCCTTCTGCCTATGATACTGCTTGGGTTGCAAGGAT
ACCTGCTGTGGATGGCTCTAATGGCCCTCAGTTTCCTCAGACGCTTCAATGGATTCTGCAAAACCAGTTGA
GCGATGGTTCCTGGGGCGAAGAATTGTGCTTCTTGACATATGACCGTGCTCTGGCTACTCTTGCCTGTGTT
ATTACGCTCTCTCTGTGGAATACAGGGGAAGAACAAGTGAACAAAGGTGTTGAATTTATAAAGAAACATGC
TGAGCGAATGGAAGGAGAAGCGGACAATCACCGGCCAAGTGGATTCGAAATTGTTTTCATTTCAATGCTCA
ACGAAGCTAAGACTCTAGGATTGGATCTACCTTATGATCTGCCTTTCTTTAAGCAGATTAATGAAATGCGC
GAAACTAAGCTTAAAAAGATTCCTTTGAATGTTGTCCACGCCATTCATACAACAATTCTGTACTCGTTGGA
GGGGTTGCAGGAAATAATAGACTGGGACAAGATAATGAAGCTTCAATCCAAGGATGGATCCTTCCTTAGCT
CCCCTGCTTCCACAGCAGCCGTGTTCATGCGCACAGGAGACAAAAAGTGCTTAGACTTTTTGAGCTTCGTT
CTCAACAAATTTCAAGATCATGCTCCTTGTCACTATCCTCTTGATCTCTTCGAGCGTCTATGGGCTGTTGA
CACTTGCCAACGCCTGGGAATCGATCGCCATTTCAAAGACGAGATCAAAGATACTCTAGATTATGTCTACA
GCTATTGGAACGATAGAGGCATCGGATGGGCAAGAGAAAATGCAGTTGGTGATATTGATGATACTGTCATG
GGCCTTCGCCTCCTCAGACTTAGCGGATACAATGTGTCATCAGATGCACTGAAAACTTTCAGAGACGAGAA
TGGAGAATTCTTTTGCTTCATGGGTCAAACACAGAGAGGAGTTACAGACATGTTAAACGTTCATCGCGGCT
CACAAGTTGCATATCCAGGCGAAACAGTCATGGAAGAAGCAAAACAGTGCACGCTCAGATATTTGACCAGC
GCTCTTGAAAATGTCGGCGCTTTTGACAAATGGGCCCTGAAAAAGGATCTTCAAGGAGAGGTGGAATACGT
GCTTAGATATCCATGGCATAGAAGCCTGCCAAGACTGGAGGCAAGGAACTACGTAGAACAGTACGGTGCAA
ATGATGTGTGGCTTGGAAAGTCCATGTACCTGATGTATTACGTAAGCAATCAGAAATATTTAGAGCTGGCC
CAAATAGACTTCAACAAGGTTCAAGCGGTACACCAAAAGGAAATTCAGGAGCTTCAGAGGTGGTGGAAATC
ATCTGGTTTCACAAAGCTCAATTTCACGCCTGAACGTGTGGTTGAAATCTATTTTGGAGTGGCGGCGACCA
TGTTTGAGCCAGAGCTAGCTACACTAAGAGCTGTTTACACCAAAACTTCAATCTTTTCTGTCATTCTCAGC
GATCTCTATGAAGCGCAGGGATCAACTCACAACATGGCGTTGTTCTCTGAAGCAGTCAAAAGATGGGATAT
GTCAATGGTTGACCGCATGCCAGAAGAAATGAAAATCTGCTTCAAGGGCTTGTATGAAACAGTCAATGAAT
TCGCAGAGGAAGGAAGGAAGAGACAAGGGCGTGACGTATTGCCTTACCTCCGAAACCTGTGGGAGGTTCAG
TTGGAATCCTACACCAAGGAAGCAAAAATGGCTCAGGCCAAATACATGCCTTCCTTCCAAGAGTACATGGA
GAATGCCAAAGTATCAATGGGACTAGCCACCATAGTTTTGACCTCAGCTCTCTTTACAGGAGATCTTCTAT
CTGATGAAACACTTTCCAAAATGGGCTATGATTCCAAATTTGTCCAATTAATGGTTTCAACAGGGCGTTTG
GTTTGTGATACAAAAACTTACCAGGAGAAACAAGAGAGAGGTCTCCCTTCAGCTGTACAATGCTACATGAA

AGACCATCCTGAAATCTCAGAAAAAGAAGCTTTAGACCATATCTACTCAGTACTGCAGAATTCCCTGCTGG
AATTGAATTGGGAACTTGTTAACAACAGAGAAATGCCAGAGACTTGTAGAAGCCTGGTTTTTAACACAGCA
AGAATAATGCAGCTATTTTACATGGAGAGAGATGGTTTTATTATGTCCCATCTAGAAATGCAACAACTTGT
GAGAAAGTGCCTTTTCGAGCCAGTGGCAGTATAGAAATCAAACTAGACAAAGACCAACATTTTTTTTCTG
TCTTAGGCATACAATATCCATGATGGGTCTCTCAGAAACTACCTTTTTTCTCGCAAAATTCTTTCTTTTTT
CTAGGATACTTCAGACTTCCATCAATTCGCTGCAAAAATAAATGCAAGTTGTATAGATTAGGTGAATCTAC
TCTTGCCTGATATGACCCTCACTTGGAGAGGGATTTGGGAAGACTTTACACTCACATATGTGGGTTTTATT
TACGCTAGTTATGTAAATTTCCATACTAGTTTGAATAATTGTTATACATATTCACATCAAAATATTTTGGT
GAACATTTTTATATAAGTCCAAGCTCT

>TpdiTPS3

MAQSLISSANSCIAKTRLSDVNGRTELNRSRPLSFSFSSSKLVKRQWNKQSAGLVLACLGDSRRSSSPAAT
GAATASTSVKREYPPAVWNDDVINSLISTYKSADVAEQEKRSETLIAEIKGMFKSMGDGETNPSAYDTAWV
ARIPAVDGSNGPQFPQTLQWILQNQLSDGSWGEELCFLTYDRALATLACVITLSLWNTGEEQVNKGVEFIK
KHAERMEGEADNHRPSGFEIVFISMLNEAKTLGLDLPYDLPFFKQINEMRETKLKKIPLNVVHAIHTTILY
SLEGLQEIIDWDKIMKLQSKDGSFLSSPASTAAVFMRTGDKKCLDFLSFVLNKFQDHAPCHYPLDLFERLW
AVDTCQRLGIDRHFKDEIKDTLDYVYSYWNDRGIGWARENAVGDIDDTVMGLRLLRLSGYNVSSDALKTFR
DENGEFFCFMGQTQRGVTDMLNVHRGSQVAYPGETVMEEAKQCTLRYLTSALENVGAFDKWALKKDLQGEV
EYVLRYPWHRSLPRLEARNYVEQYGANDVWLGKSMYLMYYVSNQKYLELAQIDFNKVQAVHQKEIQELQRW
WKSSGFTKLNFTPERVVEIYFGVAATMFEPELATLRAVYTKTSIFSVILSDLYEAQGSTHNMALFSEAVKR
WDMSMVDRMPEEMKICFKGLYETVNEFAEEGRKRQGRDVLPYLRNLWEVQLESYTKEAKMAQAKYMPSFQE
YMENAKVSMGLATIVLTSALFTGDLLSDETLSKMGYDSKFVQLMVSTGRLVCDTKTYQEKQERGLPSAVQC
YMKDHPEISEKEALDHIYSVLQNSLLELNWELVNNREMPETCRSLVFNTARIMQLFYMERDGFIMSHLEMQ
QLVRKCLFEPVAV

Figure A.4. Nucleotide and predicted amino acid sequences of three *Thuja plicata* diterpene synthases described in Chapter3.

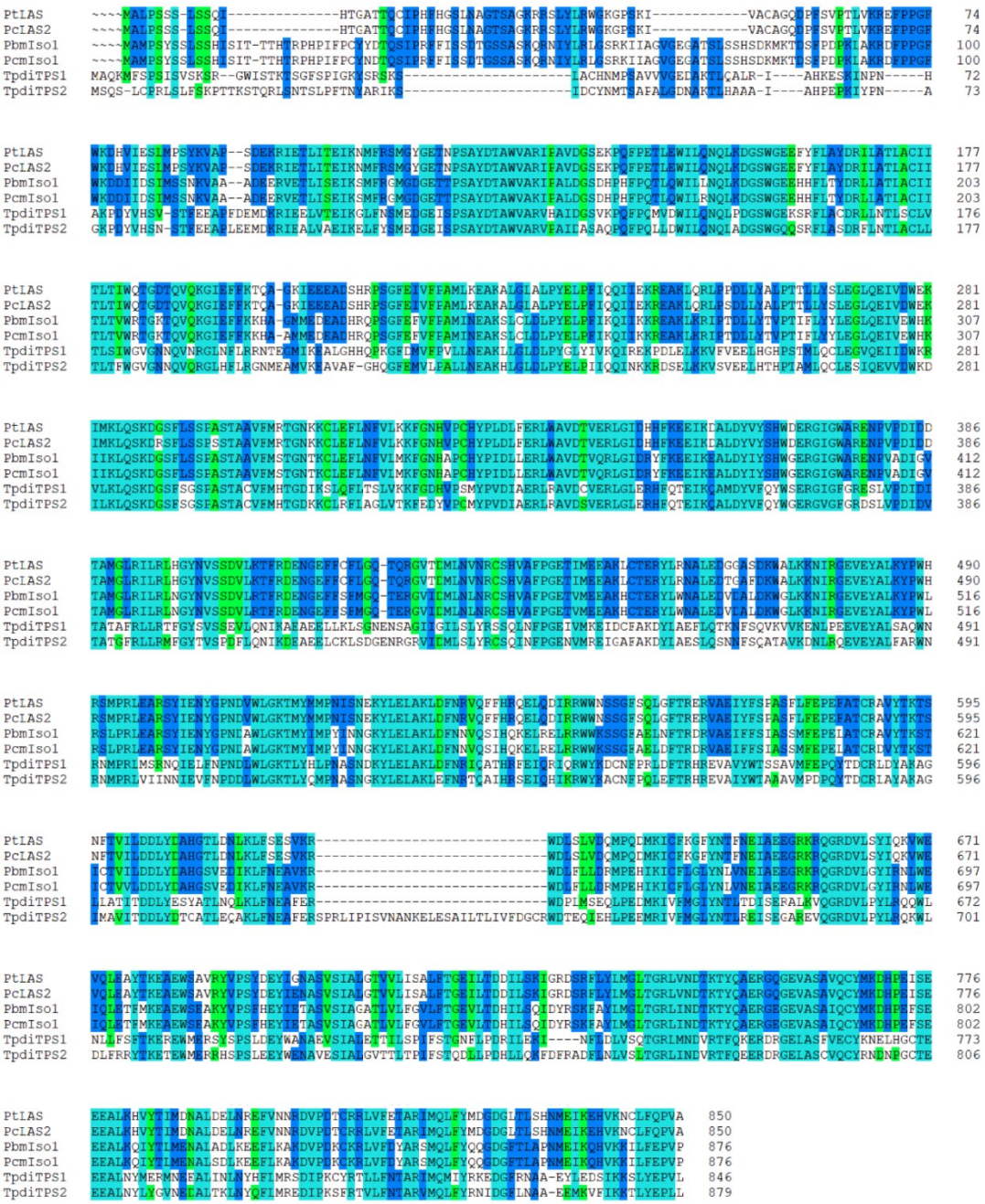


Figure A.5. Amino acid alignment of *Thuja plicata* monofunctional diTPS1 (TpdiTPS1) and diTPS2 (TpdiTPS2) with *Pinus taeda* bifunctional levopimaradiene synthase (PtLAS, Q50EK2), *Pinus contorta* bifunctional levopimaradiene/abietadiene synthase (PcLAS2, JQ240311), *Pinus banksiana* (PbmIso1, JQ240313), and *Pinus contorta* (PcmIso1, JQ240314) monofunctional class I isopimaradiene synthase1.

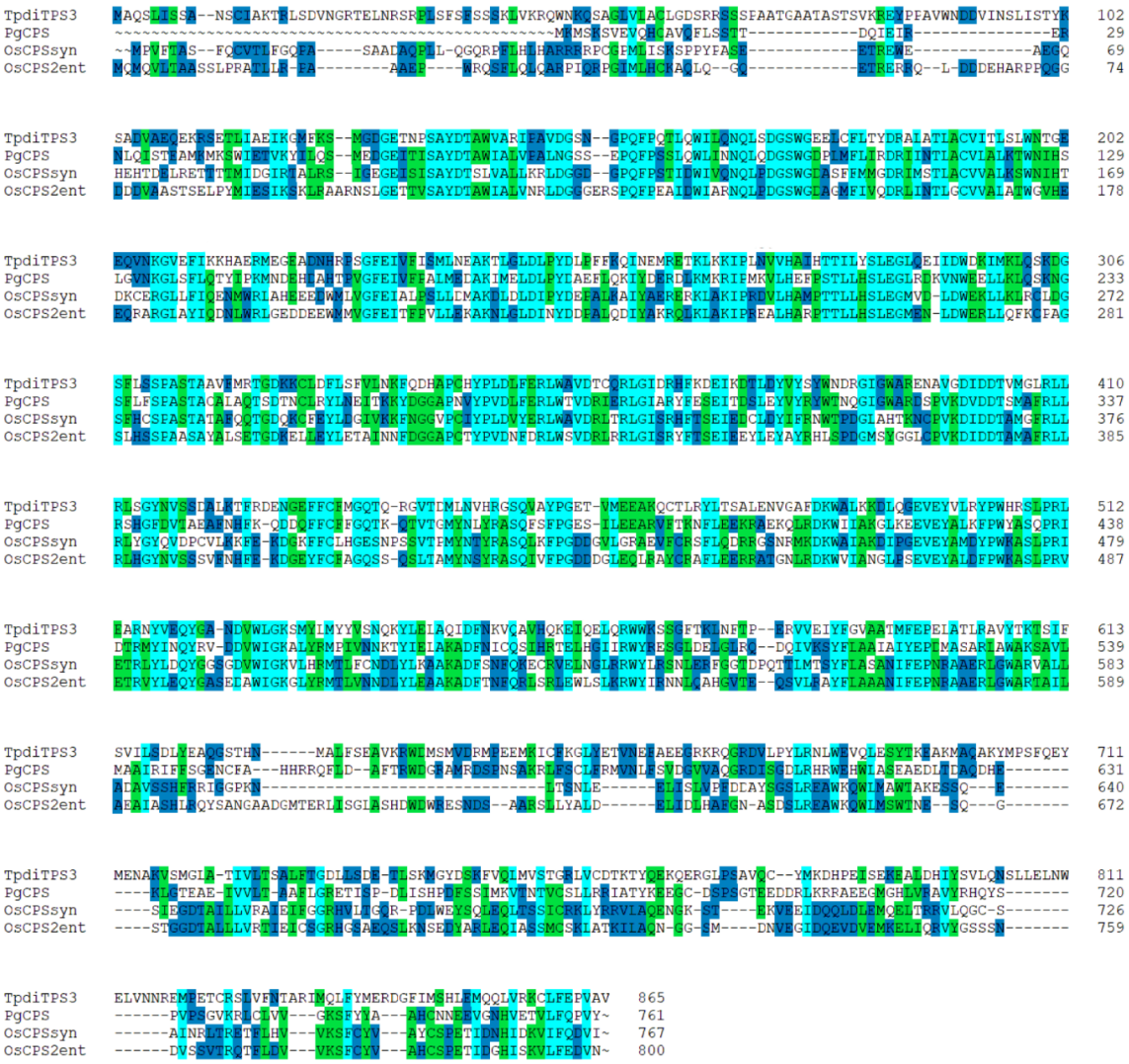


Figure A.6. Amino acid alignment of *Thuja plicata* monofunctional diTPS3 (TpdiTPS3) with *Picea glauca* monofunctional *ent*-copalyl diphosphate synthase (PgCPS, GU045755), *Oryza sativa syn*-copalyl diphosphate synthase ($OsCPS_{syn}$, AY530101) and *Oryza sativa ent*-copalyl diphosphate synthase ($OsCPS2_{ent}$, AY602991).

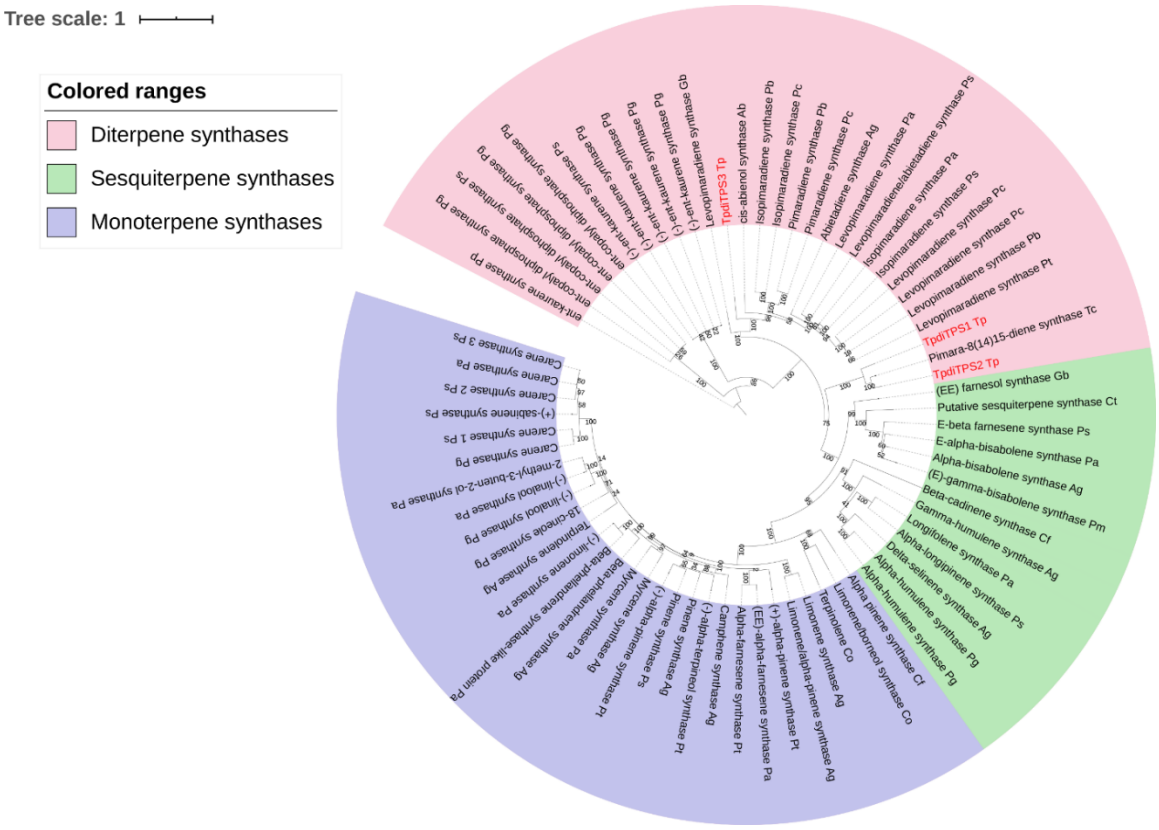


Figure A.7. Phylogenetic relationships of conifer terpene synthases. The number of bootstrap replications was 100. The abbreviation of species are as follows: *Physcomitrella patens* (Pp) was used to root the tree. *Picea glauca* (Pg), *Picea sitchensis* (Ps), *Picea abies* (Pa), *Physcomitrella patens* (Pp), *Abies grandies* (Ag), *Pinus taeda* (Pt), *Pinus contorta* (Pc), *Pinus banksiana* (Pb), *Taiwania cryptomerioides* (Tc), *Abies balsamea* (Ab), *Chamaecyparis obtuse* (Co), *Chamaecyparis formosensis* (Cf), *Pseudotsuga menziesii* (Pm), *Cycas taitangensis* (Ct).

>TpDOX1

AAAGAGGAAGAAGTTCTACATAAACCGTTCTGTTTGTTGAGATAAACACAGATTCAGCACATATCAAAGAG
AAATAATTAACACTCAGTAAAGAGGAAGAAGTTCTACATAAACCGTTCTGTTTGTTGAGATAAACACAGAA
TTAACACTCAGTAAAGAGGAAGAAATTGTTAGCAATGGAGTTCGAGGAAGGTTTTATTCAGAAGGAAGAGC
ATCGCCCCAATCCGAAACGTTCCCCCGAGGACGTGGACAATAATCTGAAAGTGCCGATTATAGACATGGAA
AAGTGCTCGGTGGAGGAAGTGGGGAAGGCGTGCAAAGAGTGGGGATTCTTCTACCTCGTTAACCACGGTGT
TCCTCATCATCTCATCCGCCGCCTCGAGTCCGCCGCCCCCGATTTCTTCTCCTTGCCCCTCGAACAAAAGA
GACGAGTTTCCAGAGATGCTGAAAACCCGTTGGGCTACTTCAACTCCGAGCGTACCAAAAACGTTAGGGAT
TGGAAAGAGGTCTTCGACTTTGCCCCGCGCGGGAAGATCGAGCTGCCCATCTCTGTTGATCCGCACGACAA
CGCTACGGTGACCGCCAGTAACCGGTGGGCGGAGGAGCTTCCGGAGTTCAGAGAAGCTTGCCTGGCGTACG
CGGAGGCTGCGAGGGAGTTGCTGGATAGGGTGCTCCACTTGATTGCACTCAGCCTCGGTTTACCTGAGAAC
CGCTTTGACAAGTACTTCGAAGATTGCATGAGCATTGCGCGGCTGAATAGTTACCCGGAATGCCCGAACCC
GGAGCTGGCGCTGGGCATAGGGAGACACAAGGACGTTGGAGCGCTTACTCTTCTTTACCAGGATGACGTGG
GTGGACTGGAGGTGAAGTGCAAGAGCAATGGGGAGTGGGTACCTGCTCCCCCCATCCCCTTCTCCTTCGTC
CTCAATGTTGCCGACTGCGTCCAGGTGTTGACGAATGATAGGTACGAGAGCGTAGAGCACAGAGCAGTGGT
AAATGACACACGAAGGAGGCTGTCAATTCCGCTATTCTTGATTCCTTCACACTCTGCGATGATAAAGCCGA
TGGAGGAATTGGTGAGCGAGGAGAATCCTCCCAAATTCAGGGAGTTCAGCTGGGGGAAGTTTACAAAGCAG

AGGATGCGTGAAGGTCTCAAAAACTTGGGGGTTGCTGACATACAAATCGACAACTTCAGAATCTGATGGAG
TGGAGTGATTATCTCATGGAATGGAAATGTGCTAAGCATGAATAATTACTACAAATAAAAATACTGCTATG
TATAATAAGCCATCTTCTGTTTGGAAAGATTGCAATTCTACCGTATCTTTTTATTATTATATCTATTCATA
TTAGTTGATTTTCTACTGTGTAGTGGTCCGTAGGGAATGATAACTAGCTTATCTGCCCTTATTGTGGCTTG
TTTCTCGCTTGGGTCAATGAATTTTTGATCTATGGGAATGAAAATGTGTTTTCAATGTGGAGTGGTGTGGA
TCATAGGTTGTGTTGGACTATTATCAAAATTTGGGAAGAAAGGTCTTTTAAAGAATATTCTTGTTTCACAT
GTGTCGATCAAGGGTGTACGTTTCTTTAGGTTTGATGAAACTTGTGATTGGTGTGTTGTGTCAAGATTCAA
CTTTTCTTTGTATGGGATTGGGACCTTG

>TpDOX1

MEFEEGFIQKEEHRPNPKRSPEDVDNNLKVPIIDMEKCSVEEVGKACKEWGFFYLVNHGVPHHLIRRLESA
APDFFSLPLEQKRRVSRDAENPLGYFNSERTKNVRDWKEVFDFAPRGKIELPISVDPHDNATVTASNRWAE
ELPEFREACLAYAEAARELLDRVLHLIALSLGLPENRFDKYFEDCMSIARLNSYPECPNPELALGIGRHKD
VGALTLLYQDDVGGLEVKCKSNGEWVPAPPIPFSFVLNVADCVQVLTNDRYESVEHRAVVNDTRRRLSIPL
FLIPSHSAMIKPMEELVSEENPPKFREFSWGKFTKQRMREGLKNLGVADIQIDNFRI

>TpDOX2

TCATTTTATCTGAACAGATGTAAATCAAATCAAAGCACCACACAGCGAGAAATAAGCAATTAGTAACCATG
GGAATTGGTACGCACCAAAACAGCCAAACTCTCATGTTCGGAGAGAGTTTTATCCAGAGCGAAGAGCATCG
TCCCAAGCATTTCTCCGCTTCTCCTGACGATAATGTTGAAATCCCCATTATAGACCTGGAAAAGAGCTCCC
CAGAGGATGTAGAAAAGGCGTGCAGAGACTGGGGTTTCTTCCACCTGGTCAATCATGGCTTTCCAGAGCAT
CTTCTGCACCGTCTTAAATCTGCCGCAGCAGATTTCTTCTCGTTGCCTGTGGAAGAAAAGAGACGAGTATC
CAGGGATGCTCAAAACCCCTTTGGCTACTTTGATTCGGAGCTAACAAAAAACGTGAGGGATTGGAAAGAGG
TATTCGACTTTGTGCCGCGGGAGGAAATCCGATTGCCCACTTCTGTTGATCCGACTGATACTGCTACGGAG
ATCTTCACAAACCGCTGGCCGGATGGCCACCCGGAATTCAGAGAAGCTTGCCTGACATATGTAGCGGCTGC
AAGAGAGTTGTCATTTAGGTTGTTACAGTTGATAGCACGAAGCCTTGGTTTACCGGAAAACCGCTTCAATG
AATATTTCAAAGATGATATAAGCAATGTGCGGATAAATAGTTACCCGGAGTGCCCGGCCCCAGAGCTCGCA
CTGGGATTGGGCAGGCACAAGGATGGCGGCGCAGTCACCCTTCTTTACCAAGACGAGGTTGGTGGTCTGCA
GGTCAAACGCAAGGACAACGGTCAATGGCTTCCCGCTCAGCCCCTGCCCAACTCCTTCGTTGTCAATGTTG
GAGATTGCATTCAGGTGTGGAGCAATGATAGGTATGAGAGCGTAGAGCACAGAGTGGTGGTGAATGATAGT
CGAAGACGTCTGTCCATTCCATTTTTCCTGAGACCTTCCCAGTATGTGATGATGAAACCTCTGGATGAGTT
AGTGGGTGAAGAGAATCCTCCCAAGTACAGCGAATTCAACTGGGGGAAGTTTTTCAAGAGGAGGAAGGATG
GCAATTTTAAAAATTTAGGGGTTGAGAATCAACAAATATATCATTTTAGAATCTGAGTGGTATGAGA

>TpDOX2

MGIGTHQNSQTLMFGESFIQSEEHRPKHFSASPDDNVEIPIIDLEKSSPEDVEKACRDWGFFHLVNHGFPE
HLLHRLKSAAADFFSLPVEEKRRVSRDAQNPFGYFDSELTKNVRDWKEVFDFVPREEIRLPTSVDPTDTAT
EIFTNRWPDGHPEFREACLTYVAAARELSFRLLQLIARSLGLPENRFNEYFKDDISNVRINSYPECPAPEL
ALGLGRHKDGGAVTLLYQDEVGGLQVKRKDNGQWLPAQPLPNSFVVNVGDCIQVWSNDRYESVEHRVVVND
SRRRLSIPFFLRPSQYVMMKPLDELVGEENPPKYSEFNWGKFFKRRKDGNFKNLGVENQQIYHFRI

>TpDOX3

AGCACAACAACATAACTGTAATAATGGGAATTGGGGCGGCCGAGAACACCGAAGCTCTGAAGTTCGAAGAG
AGCTTTATTCAGAGCGAAGAGCATCGTCCCAAGCATTTCTCCTCTCCTCCTTCCGATAATGTCAAAATACC
GGTGATAGACTTGGAAAAGAGCTCTGCAGAGGATGTGGAGAAGGCGTGCAGAGAGTGGGGTTTCTTCCACC
TCGTCAATCATGGCTTCCCTGATCATCTTCTGCACCGCCTCAAATCCGTTGCAGCCGACTTCTTCTCGTTG
CCTCTCGAAGAAAAGAGACGAGTATCAAGGGATGCAGAAAACTCTTTGGGCTACTTCAATTCGGAGCTTAC
AAAAAACGTGAGGGATTGGAAAGAGGTGTTCGACTTTGTGCCGCGGGAGGAAATCCAATTGCCTACATCTG
TTGATCCGGACAACCATGCTACCTACACCATCAAAAACCAATGGCCGGATGGCCATCCGGATTTCAGAGAA
GCTTGCCTGACCTATGCAGAGGCTGCGAGAGAATTAGTGTTTAAGTTGCTAGAATTAATAGCACGGAGCCT

TGGCTTACCTGGAAACCGGCTGGATGAGTATTTCAAAGATGATATAACTGCTATGCGGTTGAATTGTTACC
CAGAATGCCCAAATCCCGAGCTAGCACTGGGCGTGGGAAGGCACAAAGACGGCGCTGCACTTACCCTTCTT
TACCAAGACGAAGTGGGTGGATTGGAAGTGAAGTGCAAGGAGAGTGGTCAGTGGGTGCCCGCTCAACCGAT
GCCCAACTCTTTCGTTGTCAATGTTGGAGACTGCATACAGGTGTGGAGCAATGACAGGTATGAGAGTGTTG
AGCACAGAGTGGTGGTGAATGATAGCCGAAGACGCTTCTCCATTCCGTATTTCCTTAGTCCTTCCCATTAT
GTGATTATAAAGCCTCTGGAGGAGTTGGTGAATGAGGAAAATCCTCCCAAGTATAGGGAATATAATTGGGG
GAAGTTTTCTAAGCGTAGGAGGGATAGTAATTTTAAGAACTTGGGAGAAGAGAATCTTCAAATATACCACT
TCAGAATCTGACTGATATGATATTAGCGGTGTGGGATTGCGTACTCCCAAGGATAAAAAATGTCTTAAACA
GAACCTGTTACTAATGAATAAATCTCATAAATGTGCATGTTGTAAATATGGTAAAATAAACTCTCCATATG
TATAAGTCATGGTGTGTAGAAAGGTTGCATCTCTACAGTATATTAAAATAAACTCTCCATATGTATAAGTC
ATGGTGTGTAGAAAGGTTGCATCTCTACAATATATTTGGTAATGGTATTATATATATCATATCAACTCTCA
AATGCTGATTCTTCCTTGTAACGCTTCATGAAAGCTTAAAATATT

>TpDOX3

MGIGAAENTEALKFEESFIQSEEHRPKHFSSPPSDNVKIPVIDLEKSSAEDVEKACREWGFFHLVNHGFPD
HLLHRLKSVAADFFSLPLEEKRRVSRDAENSLGYFNSELTKNVRDWKEVFDFVPREEIQLPTSVDPDNHAT
YTIKNQWPDGHPDFREACLTYAEAARELVFKLLELIARSLGLPGNRLDEYFKDDITAMRLNCYPECPNPEL
ALGVGRHKDGAALTLLYQDEVGGLEVKCKESGQWVPAQPMPNSFVVNVGDCIQVWSNDRYESVEHRVVVND
SRRRFSIPYFLSPSHYVIIKPLEELVNEENPPKYREYNWGKFSKRRRDSNFKNLGEENLQIYHFRI

Figure A.8. Nucleotide and predicted amino acid sequences of three *Thuja plicata* dioxygenases described in Chapter4.

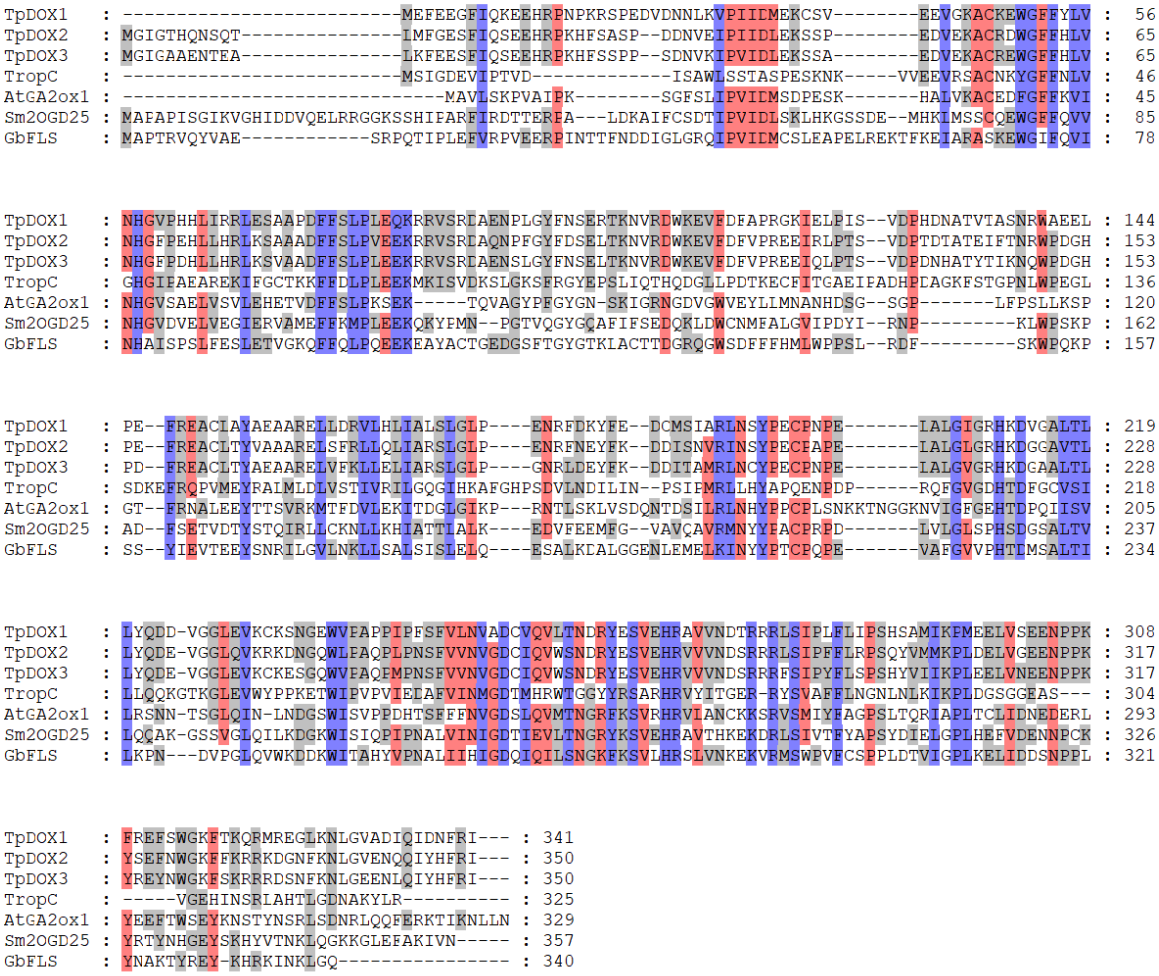


Figure A.9. Amino acid alignment of *Thuja plicata* TpDOX1, TpDOX2 and TpDOX3 with *Talaromyces stipitatus* Fe(II)/2-oxoglutarate-dependent dioxygenase (TropC,), *Arabidopsis thaliana* gibberellin 2-β-dioxygenase (AtGA2ox1), *Salvia miltiorrhiza* Fe(II)/2-oxoglutarate-dependent dioxygenase (Sm2OGD25), and *Ginkgo biloba* flavonol synthase (GbFLS).

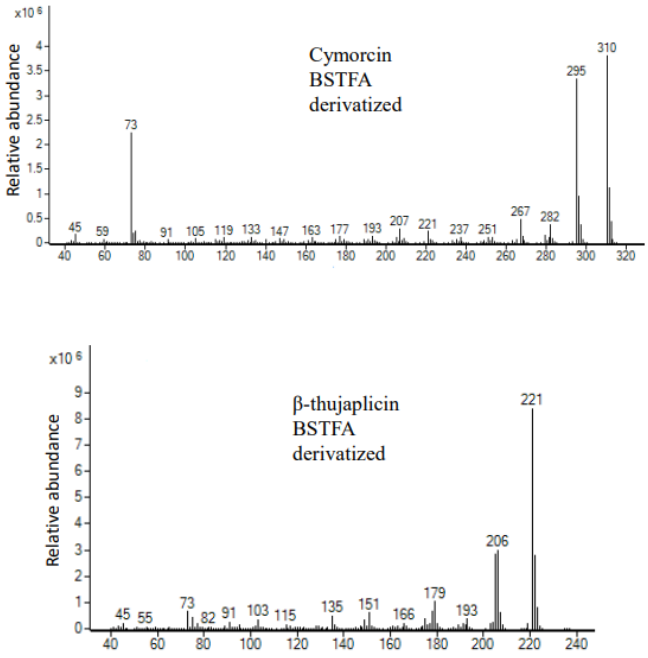


Figure A.10. Mass spectrum of derivatized Cymorcin (A) and derivatized β-thujaplicin (B).

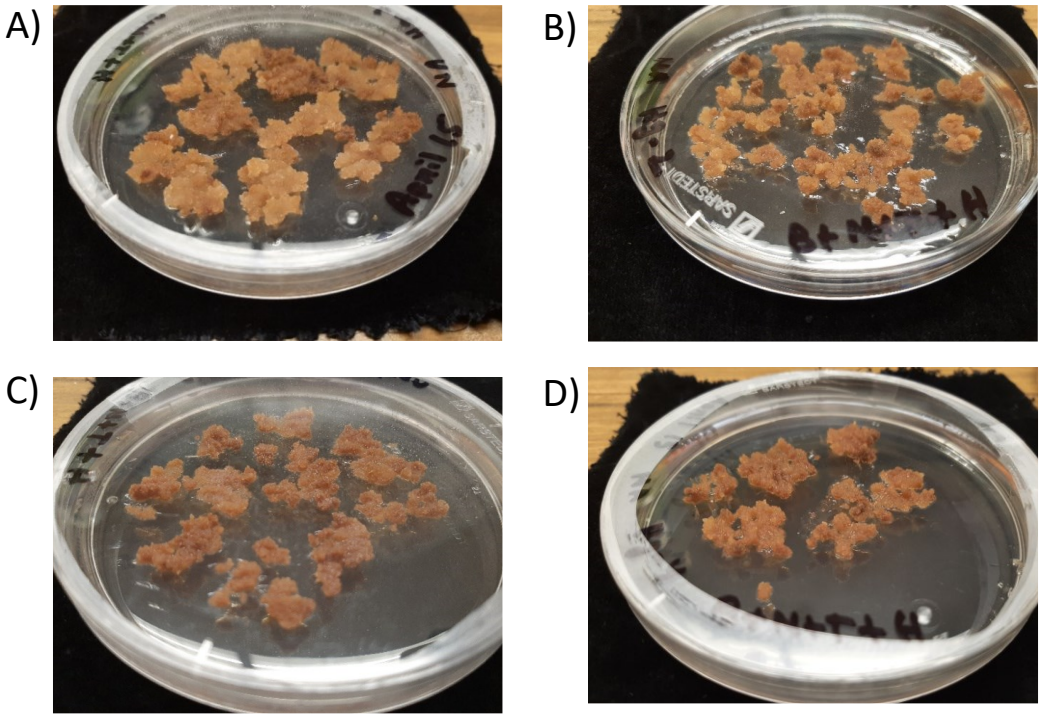


Figure A.11. Hygromycin resistance test in *C. lusitanica* calli. A) 1 µg/mL, B) 2 µg/mL, C) 3 µg/mL, D) 4 µg/mL.

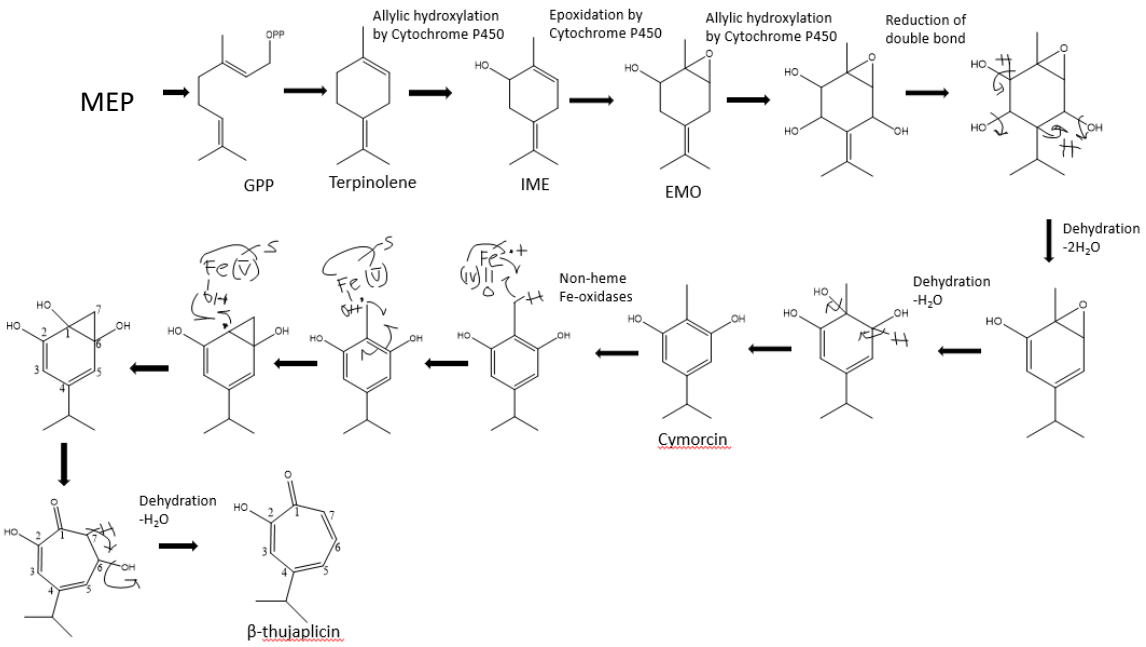


Figure A.12. Proposed β-thujaplicin biosynthesis pathway (Priyadarshini Balaraman).

Stumped Block				Unstumped Block			
Fir Birch Plot1	Fir Plot16	Cedar Birch Plot17	Larch Plot32	Cedar Plot33	Pine Cedar Plot48	Pine Birch Plot49	Fir Birch Plot64
Fir Cedar Plot2	Pine Birch Plot15	Pine Cedar Plot18	Birch Plot31	Pine Birch Plot34	Fir Pine Plot47	Pine Plot50	Spruce Plot63
Birch Plot3	Fir Pine Plot14	Pine Plot19	Fir Plot30	Cedar Birch Plot35	Cedar Plot46	Fir Cedar Plot51	Pine Cedar Plot62
Cedar Plot4	Cedar Birch Plot13	Pine Cedar Plot20	Fir Pine Plot29	Fir Plot36	Fir Birch Plot45	Birch Plot52	Fir Plot61
Spruce Plot5	Cedar Plot12	Fir Birch Plot21	Pine Birch Plot28	Fir Cedar Plot37	Birch Plot44	Fir Birch Plot53	Pine Plot60
Fir Cedar Plot6	Birch Plot11	Pine Cedar Plot22	Fir Plot27	Pine Cedar Plot38	Cedar Birch Plot43	Pine Plot54	Fir Pine Plot59
Cedar Birch Plot7	Pine Plot10	Pine Birch Plot23	Fir Birch Plot26	Fir Plot39	Fir Cedar Plot42	Cedar Birch Plot55	Birch Plot58
Fir Pine Plot8	Cedar Plot9	Fir Cedar Plot24	Pine Plot25	Fir Pine Plot40	Larch Plot41	Cedar Plot56	Pine Birch Plot57

Figure A.13. Experimental design of the Skimikin trial that has 32 plots stumped (on the left) and 32 plots unstumped (on the right) over a continuous area. Each plot measures 0.04 ha. In this study, *T. plicata* wood cores collected from plots (marked with red squares) are colored with yellow (Pure Cedar) and blue (Cedar growing with fir). Total 52 tress were collected randomly from 4 stumped and 6 unstumped plots. In stumped plots, monoculture cedar trees were collected from

plot 4 (5 trees), 12 (5 trees) and 9 (5 trees) whereas cedar trees growing with fir were collected from plot 24 (7 trees). In unstumped plots, monoculture cedar trees were collected from plot 33 (5 trees), 46 (5 trees) and 56 (5 trees) whereas cedar trees growing with fir were collected from plot 37 (5 trees), 42 (5 trees) and 51 (5 trees).

Appendix B. Supplementary tables

Table B.1. List of proteins used in phylogenetic analysis in Chapter2.

Gene ID	Function	Accession No	Species
PgTPS-3car	3-carene synthase	C7ASI9	*Picea glauca*
PsTPS-3car1	3-carene synthase1	F1CKI6	*Picea sitchensis*
PaTPS-3car	3-carene synthase	Q84SM8	*Picea abies*
PsTPS-sab	(+)-sabinene synthase	F1CKJ1	*Picea sitchensis*
PaTPS-Lin	(-)-linalool synthase	AAS47693.1	*Picea abies*
PgTPS-Lin	(-)-linalool synthase	ADZ45500.1	*Picea glauca*
Aggteo	terpinolene synthase	Q9M7D0	*Abies grandis*
Aggmyr	myrcene synthase	O24474	*Abies grandis*
PaTPS-Myr	myrcene synthase	AAS47696.1	*Picea abies*
PaTPS-Lim	(-)-limonene synthase	AAS47694.1	*Picea abies*
Agg-pin1	pinene synthase	O24475	*Abies grandis*
Agg-lim1	limonene synthase	O22340	*Abies grandis*
PaTPS-Far	E, E-alpha-farnesene synthase	AAS47697.1	*Picea abies*
Cf-Pin	α-pinene synthase	EU099434.1	*Chamaecyparis formosensis*
TpSS1	sabinene synthase	KC767281.1	*Thuja plicata*
TpTS1	terpinolene synthase	OM324392	*Thuja plicata*
TpCS1	3-carene synthase	OM324393	*Thuja plicata*
PaTPS-Lon	longifolene synthase	Q675L0	*Picea abies*
PsTPS-Lonp	alpha-longipinene synthase	ADZ45516.1	*Picea sitchensis*
Agfdsel1	delta-selinene synthase	O64404	*Abies grandis*
PgTPS-Hum	alpha-humulene synthase	ADZ45513.1	*Picea glauca*
Agfghum	gamma-humulene synthase	O64405	*Abies grandis*
CfCad	beta-cadinene synthase	AFJ23663.1	*Chamaecyparis formosensis*
AgfEabis	alpha-bisabolene synthase	O81086	*Abies grandis*
PmeTPS3	(E)-gamma-bisabolene synthase	Q4QSN4	*Pseudotsuga menziesii*
TpdiTPS1	monofunctional sandaracopimaradiene synthase	MT468207	*Thuja plicata*
TpdiTPS2	monofunctional levopimaradiene synthase	MT468208	*Thuja plicata*
PbmPIM1	monofunctional pimaradiene synthase	M4HY08	*Pinus banksiana*
PcmPIM1	monofunctional pimaradiene synthase	M4HYC8	*Pinus contorta*
PbmIso1	monofunctional isopimaradiene synthase	M4HXW5	*Pinus banksiana*
PcmIso1	monofunctional isopimaradiene synthase	M4HYP3	*Pinus contorta*
PaTPS-LAS	bifunctional levopimaradiene synthase	Q675L4	*Picea abies*
PsTPS-LAS	levopimaradiene/abietadiene synthase	ADZ45517.1	*Picea sitchensis*
PcLAS1	bifunctional levopimaradiene synthase	M4HY05	*Pinus contorta*
AgAS	bifunctional abietadiene synthase	Q38710	*Abies grandis*
CfCPS1	copalyl diphosphate synthase	QWV53997.1	*Chamaecyparis formosensis*
TpdiTPS3	normal-copalyl diphosphate synthase	MT468209	*Thuja plicata*
GbTPS-Lev	bifunctional levopimaradiene synthase	Q947C4	*Ginkgo biloba*
PgCPS	ent-copalyl diphosphate synthase	JAI17637.1	*Picea glauca*
PpCPS/KS	ent-kaurene synthase	BAF61135.1	*Physcomitrella patens*

Table B.2. Primers used for nested PCR and SLIC-based cloning into expression plasmid in Chapter2.

GENE ID	PRIMER SET	FORWARD PRIMER (5'-3')	REVERSE PRIMER (5'-3')
TpTS1	Set1	TGGCTCTTTCCTCTGCTTTT	TCAAGTAAAGAGAGACTTATGTGATTG
	Set2	TACTTCCAATCCAATGCATTAATAGTAAA TCAGGCCCAA	TTATCCACTTCCAATGTTATTACATAGAT ACAGGCTCAAT
TpCS1	Set1	TGAGTTGCATTACGCCTTTG	GAAAATCATTATTTATAAACACACACAA
	Set2	TACTTCCAATCCAATGCACTTCCAATCA ATATGGCGTT	TTATCCACTTCCAATGTTATCACATAGG AACTGGTTCAA
TpES1	Set1	AACTTGTAAATTTGAGTTGTAGTCCA	CAAGTATGGCAACCAAGAGTCA
	Set2	TACTTCCAATCCAATGCA ATGTCTAATTTGAAGGGAGA	TTATCCACTTCCAATGTTA TTAAACTTTAACTAGATCAA

Table B.3. Kovats Retention indices of identified monoterpene and sesquiterpene compounds in Chapter2.

GENE NAME	COMPOUND NAME	RETENTION INDEX (R.I)	LITERATURE (R.I)
TpTS1	Terpinolene	1092.22	1088
TpCS1	3-carene	1013.128	1011
TpES1	Elemol	1565.625	1549

Table B.4. Primers used for nested PCR and SLIC-based cloning into expression plasmid in Chapter3.

GENE ID	PRIMER SET	FORWARD PRIMER (5'-3')	REVERSE PRIMER (5'-3')
TpdiTPS1	Set1	TCAAAATGGCTCAGAAGATGT	CCATGATTGTAATGATGCCTTTT
	Set2	TACTTCCAATCCAATGCAAAGACTCT TCAGGCTTTGAG	TTATCCACTTCCAATGTTATTACAGCACTG GTTCGTACA
TpdiTPS2	Set1	TACTCCCCATGGGATTCAAA	CACAGTTCATCCAATAGAGCAA
	Set2	TACTTCCAATCCAATGCAGCGAAAACAC TTCACGCTG	TTATCCACTTCCAATGTTATTAGAGCAGG GGTTCATAGA
TpdiTPS3	Set1	CTGGTTTGGTTTTGGCTTGT	CCAAATCCCTCTCCAAGTGA
	Set2	TACTTCCAATCCAATGCAGCTGCCACTG CTTCTACATCG	TTATCCACTTCCAATGTTACTATACTGCCA CTGGCTCGA

Table B.5. Kovats Retention indices of identified diterpene compounds in Chapter3.

GENE NAME	COMPOUND NAME	RETENTION INDEX
TpdiTPS1	sandaracopimaradiene	1986
	syn-stemod-13(17)-ene	2119
	unknown diterpene	1981
TpdiTPS2	levopimaradiene	2039
TpdiTPS3	normal-copalol	2293

Table B.6. Primers used in cloning of three putative dioxygenases in pET28b expression vector in Chapter4.

GENE ID	FORWARD PRIMER (5'-3')	REVERSE PRIMER (5'-3')
TpDOX1	AGGAGATATACCATGATGGAGTTCGAGG AAGGTTTTA	GGTGGTGGTGCTCGAGGATTCTGAAGTTG TCGATTTGT
TpDOX2	AGGAGATATACCATGATGGGAATTGGTAC GC	GGTGGTGGTGCTCGAGGATTCTAAAATGAT ATATTTGTTGATT
TpDOX3	AGGAGATATACCATGATGGGAATTGGGG CGG	GGTGGTGGTGCTCGAGGATTCTGAAGTGG TATATTTGAAGAT

Table B.7. Primers used in cloning of three putative dioxygenases in pWBVec8-Ubi:USER:NOS expression vector in Chapter4.

GENE ID	FORWARD PRIMER (5'-3')	REVERSE PRIMER (5'-3')
TpDOX1	AGGAGATATACCATGATGGAGTTCGAGG AAGGTTTTA	GGTGGTGGTGCTCGAGGATTCTGAAGTTG TCGATTTGT
TpDOX2	AGGAGATATACCATGATGGGAATTGGTAC GC	GGTGGTGGTGCTCGAGGATTCTAAAATGAT ATATTTGTTGATT
TpDOX3	AGGAGATATACCATGATGGGAATTGGGG CGG	GGTGGTGGTGCTCGAGGATTCTGAAGTGG TATATTTGAAGAT

Table B.8. RTqPCR primers for validation of overexpression in C. lusitanica and C. cashmeriana calli in Chapter4.

GENE ID	FORWARD PRIMER (5'-3')	REVERSE PRIMER (5'-3')
TpDOX1	CCGCTATTCTTGATTCCTTCAC	TTGAGACCTTCACGCATCC
TpDOX2	CACAGAGTGGTGGTGAATGATAG	CCACTAACTCATCCAGAGGTTTC
TpDOX3	GAATGATAGCCGAAGACGC	TCCCTCCTACGCTTAGAAAAC
TpTS1	TCCACTGAAGAAGATGCCTTAAA	ACTACAAGTTGGAACACCATCA
16SrRNA	TCGCGCACCTTAGAAATCC	TCGCATTTCGCTACATTCTTC

Table B.9. Protein sequences used in phylogenetic analysis.

Gene ID	Function	Accession No	Species
AtGA3ox1	gibberellin 3-beta-dioxygenase	At1g15550	*Arabidopsis thaliana*
AtGA3ox2	gibberellin 3-beta-dioxygenase	At1g80340	*Arabidopsis thaliana*
AtGA3ox3	gibberellin 3 beta-hydroxylase	At4g21690	*Arabidopsis thaliana*
AtGA3ox4	gibberellin 3 beta-hydroxylase	At1g80330	*Arabidopsis thaliana*
AtGA20ox1	gibberellin 20-oxidase	At4g25420	*Arabidopsis thaliana*
AtGA20ox2	gibberellin 20-oxidase	At5g51810	*Arabidopsis thaliana*
CmGA20ox	gibberellin 20-oxidase	AAB64345	*Cucurbita maxima*
AtGA2ox1	carbon19-gibberellin 2-oxidase	At1g78440	*Arabidopsis thaliana*
AtGA2ox2	carbon19-gibberellin 2-oxidase	At1g30040	*Arabidopsis thaliana*
AtGA2ox3	carbon19-gibberellin 2-oxidase	At2g34555	*Arabidopsis thaliana*
AtGA2ox7	carbon20-gibberellin 2-oxidase	At1g50960	*Arabidopsis thaliana*
AtGA2ox8	carbon20-gibberellin 2-oxidase	At4g21200	*Arabidopsis thaliana*
OsDAO	Dioxygenase for Auxin Oxidation	NP_001053075.1	*Oryza sativa*
AtAOP1	glucosinolate biosynthesis	NM_116541	*Arabidopsis thaliana*
AtAOP2	glucosinolate biosynthesis	AF417858	*Arabidopsis thaliana*
AtF3H	flavanone 3-hydroxylae	At3g51240	*Arabidopsis thaliana*
GbF3H	flavanone 3 beta-hydroxylase	AY742228.1	*Ginkgo biloba*
GmF3H	flavanone 3 beta-hydroxylase	AY595420.1	*Glycine max*
AtFLS1	flavonol synthase	At5g08640	*Arabidopsis thaliana*

GbFLS	flavonol synthase	ACY00393.1	*Ginkgo biloba*
GbANS	anthocyanidin synthase	ACC66092.1	*Ginkgo biloba*
OsANS	anthocyanidin synthase	Y07955.1	*Oryza sativa*
PaANS	anthocyanidin synthase	AB198870.1	*Phytolacca americana*
AtACO4	ACC oxidase	At1g05010	*Arabidopsis thaliana*
AtACO3	ACC oxidase	At1g12010	*Arabidopsis thaliana*
AtACO2	ACC oxidase	At1g62380	*Arabidopsis thaliana*
AtACO5	ACC oxidase	At1g77330	*Arabidopsis thaliana*
AtACO1	ACC oxidase	At2g19590	*Arabidopsis thaliana*
SlACO1	1-aminocyclopropane-1-carboxylate oxidase	NM_001247095	*Solanum lycopersicum*
Sm2OGD25	diterpenoid biosynthesis	Not Available	*Salvia miltiorrhiza*
At2OGD	diterpenoid biosynthesis	NP_566685	*Arabidopsis thaliana*
ZmFLS/F3H	diterpenoid biosynthesis	NM_001357949	*Zea mays*
AtJOX1	jasmonic acid biosynthesis	NP_187728	*Arabidopsis thaliana*
AtJOX2	jasmonic acid biosynthesis	Q9FFF6	*Arabidopsis thaliana*
AtJOX3	jasmonic acid biosynthesis	Q9LY48	Arabidopsis thaliana
AtJOX4	jasmonic acid biosynthesis	AEC09512	Arabidopsis thaliana
AtJRG1	jasmonic acid biosynthesis	NP_191156	Arabidopsis thaliana
TropC	2-oxoglutarate/Fe(II)-dependent dioxygenase	B8M9K5	Talaromyces stipitatus
TpDOX1	Putative 2-oxoglutarate/Fe(II)-dependent dioxygenase	Not Available	Thuja plicata
TpDOX2	Putative 2-oxoglutarate/Fe(II)-dependent dioxygenase	Not Available	Thuja plicata
TpDOX3	Putative 2-oxoglutarate/Fe(II)-dependent dioxygenas	Not Available	Thuja plicata

Appendix C. Composition of buffers, media, and stock solutions

2TY medium: for 1 L

Tryptone 16 g, yeast extract 10 g and NaCl 5 g were dissolved in dH_2O.

½ MS (Murashige and Skoog) medium: for 2 L, pH 5.5-5.8

MS 4.44 g, sucrose 30 g, MES 1 g, and agar 4g/0.5L.

Gamborg B5 medium (callus culture medium): for 1 L

Micro elements 100X stock for 1 L ($CoCl_2.6H_2O$ 0.0025 g, $CuSO_4.5H_2O$ 0.0025 g, FeNaEDTA 3.67 g, H_3BO_3 0.3 g, KI 0.075 g, $MnSO_4.H_2O$ 1 g, $Na_2MoO_4.2H_2O$ 0.025 g, $ZnSO_4.7H_2O$ 0.2 g), stored at 4°C.

Vitamins 100X stock for 200 mL (myo-inositol 2 g, nicotinic acid 0.02 g, pyridoxine HCl 0.02 g, Thiamine HCl 0.2 g), stored at 4°C.

Hormones:

NAA (1-Naphthaleneacetic acid) 100 mM stock for 1 mL:

0.0186 g NAA dissolved in 1 mL DMSO, stored at -20°C.

BAP (6-Benzylaminopurine) 100 mM stock for 1 mL:

0.0225 g dissolved in DMSO; a 1 mM stock was further prepared from this 100 mM stock. stored at -20°C.

To prepare the B5 medium for 1 L, 10 mL of micro elements (100X stock), macro elements: $CaCl_2.2H_2O$ 0.15 g, KNO_3 2.5 g, $MgSO_4$ 0.12 g, $NaH_2PO_4.H_2O$ 0.15 g, 10 mL vitamins (100X stock), 20 g/L sucrose, Gelzan 2g/0.5L. The pH was adjusted to 5.5. The media was autoclaved cooled and added 0.01 µM BAP and 10 µM NAA (Added fresh).

Timentin: 100 mg/mL for 10 mL

1 g timentin dissolved in 10 mL dH_2O, filter-sterilized and stored at -20°C. Always add fresh in medium.

Acetosyringone: 1 M for 1mL

0.1962 g acetosyringone was dissolved in 1 mL DMSO and stored at -20°C.

IPTG:

200 mM stock prepared in dH_2O, filter-sterilized and stored at -20°C.

Dioxygenase protein purification buffer:

Lysis buffer: Prepared fresh and store on ice

1 mg/mL DNase, 1 mg/mL RNase, 1 mM $MgCl_2$, 1 mM PMSF, and 25 mg lysozyme were added in HisTrap binding buffer pH 7.5.

HisTrap Binding buffer

50 mM Tris-HCl, 150 mM NaCl, 5 mM imidazole and 5% v/v glycerol were added to dH_2O and the final pH was adjusted to 7.5. The solution was filter sterilized and stored at 4°C.

HisTrap Elution buffer

50 mM Tris-HCl, 150 mM NaCl, 500 mM imidazole and 5% v/v glycerol were added to dH_2O and the final pH was adjusted to 7.5. The solution was filter sterilized and stored at 4°C.

Desalting buffer

50 mM Tris-HCl, 150 mM NaCl and 5% v/v glycerol were added to dH_2O and the final pH was adjusted to 7.5. The solution was stored at 4°C.

Dioxygenase assay buffer:

Tris-HCl, pH 7.5 (1 M, 1 L)

Add 121.14 g of Tris base in 800 mL of distilled water. Adjust pH using HCl. Add distilled water until the volume is 1 L.

N-[Tris(hydroxymethyl)-methyl]-2-aminoethanesulfonic acid (TES) buffer: 200 mM, 30 mL, pH 7.5

1.38 g of TES dissolved in sterile dH_2O. The pH was adjusted with 1 N NaOH to pH 7.5. The final volume was adjusted up to 30 mL. The solution was filter sterilized and stored at 4°C.

2-oxoglutaric acid (100 mM, 2 mL):

0.029 g 2-oxoglutaric acid dissolved in sterile dH_2O (prepared fresh)

ammonium iron (II) sulfate hexahydrate [$(NH_4)_2Fe(SO_4)_2.6H_2O$] (10 mM, 10 mL):

0.039 g $(NH_4)_2Fe(SO_4)_2.6H_2O$ dissolved in sterile dH_2O (prepared fresh)

Sodium ascorbate (100 mM, 2 mL):

0.039 g sodium ascorbate dissolved in sterile dH_2O (prepared fresh)

Cymorcin (250 mM, 0.5 mL):

0.021 g cymorcin dissolved in 500 µl dimethyl sulfoxide (DMSO) and stored at -20°C.

Authentic β-thujaplicin:

0.01 g β-thujaplicin was dissolved in 1 mL ethyl acetate (EtOAc) to make a stock concentration of 10 mg/mL which was diluted 100-fold with EtOAc to make a concentration of 100 ng/µl from which 100 µl of sample was taken and derivatized with 5 µl *N,O*-Bis(trimethylsilyl)trifluoroacetamide (BSTFA), incubated for 2 h. and detected on GC-MS.

2X SDS-PAGE sample buffer (Laemmli, 30 mL):

3.75 mL of 0.5 M Tris-HCl pH 6.8, 15 mL of 50% glycerol, 0.3 mL of 1% Bromophenol blue and 3 mL of 20% SDS were added to sterile dH_2O.

1X SDS-PAGE running buffer (1 L):

100 mL of 10X Tris/Glycine/SDS buffer (purchased from Bio-Rad) was diluted with 900 mL dH_2O.

Antibiotics:

Spectinomycin: A stock solution of 50 mg/mL for 10 mL was prepared by dissolving 0.5 g spectinomycin in 10 mL sterile dH_2O. The solution was filter sterilized and stored at -20°C.

Gentamycin: A stock solution of 10 mg/mL for 10 mL was prepared by dissolving 0.1 g gentamycin in 10 mL sterile dH_2O. The solution was filter sterilized and stored at -2°C.

Kanamycin: A stock solution of 50 mg/mL for 10 mL was prepared by dissolving 0.5 g kanamycin in 10 mL sterile dH_2O. The solution was filter sterilized and stored at -20°C.

Appendix D. Codes for mixed effect model analysis

```
#find working directory
getwd()
#set working directory
setwd("C:/Users/user/Desktop/Chapter5_skimikin/skimikin")
#Install edgeR
#if (!requireNamespace("BiocManager", quietly = TRUE))
#  install.packages("BiocManager")
#BiocManager::install("edgeR")
#load all required libraries
library ("edgeR")
library ("gplots")
library(RColorBrewer)
library(goseq)
library("statmod")
library("goseq")
library("GO.db")
library("qvalue")
#Open data for analysis
data <- read.csv("June_Auguststump_unstump_cwcwfdcounts.csv", row.names=1)
head(data, 4)
nrow(data)
#Provide sample information
sampleinfo <- read.csv("SampleinfoJA.csv")
head(sampleinfo, 4)
colnames(sampleinfo)[1] <- 'FileName'
#Visualize the data and save as .tiff file
tiff(file="total number of reads.tiff", width=6, height=5, units="in", res=300)
barplot(colSums(data), ylab="Total number of reads", xlab="Sample ID", cex.names=0.5)
dev.off()
#Transform raw counts to log counts and visualize sample to sample correlation matrix
log10_data <- log10(data + 1)
cor(log10_data, method="pearson")
heatmap(as.matrix(cor(log10_data, method="pearson")), main="Clustering of Pearson correlations", scale="none")
#save sample correlation matrix as .tiff file
tiff(file="sample correlation matrix.tiff", width=6, height=5, units="in", res=300)
heatmap(as.matrix(cor(log10_data, method="pearson")), main="Clustering of Pearson correlations", scale="none")
dev.off()
#Convert data into factor
plot <- factor(sampleinfo$Plot)
treatment1 <- factor(sampleinfo$Treatment1)
treatment2 <- factor(sampleinfo$Treatment2)
time<-factor(sampleinfo$Time)
#make a DGElist object out of the data
list1 <- DGEList (counts = data)
list1
#calculate the normalization factors to adjust the effective library size relative to other libraries in the dataset
list1 <- calcNormFactors (list1)
list1
logcounts <-cpm(list1, log=TRUE)
#calculate the counts per million
cpm.list1 <- cpm(list1)
head(cpm.list1, 4)
#filter out the genes with less than 1 cpm in 6 or fewer libraries
list2 <- list1[rowSums(cpm.list1 > 1) >= 6,]
head(list2,4)
cpm.list2 <- cpm.list1[rowSums(cpm.list1 > 1) >= 6,]
head(cpm.list2, 4)
#generate a multi-dimensional scaling (MDS) plot
col.cell <- c("purple","orange")[treatment1]
#plot the MDS graph and save as .tiff file
tiff(file="Multidimentional scaling plot.tiff", width=6, height=5, units="in", res=300)
```

```
legend("topleft",fill=c("purple","orange"),legend=levels(treatment1))
plotMDS (list2, col = col.cell)
legend("topleft",fill=c("purple","orange"),legend=levels(treatment1))
dev.off()
#Set the model to use
design <- model.matrix (~0+time)
rownames(design) <- colnames(data)
#or
design <- model.matrix (~0+treatment1)
#or
design <- model.matrix (~0+treatment2)
#voom calculates weights to offset the mean variance relationship
list2 <- voom(list2, design, plot=TRUE)
cor <- duplicateCorrelation(list2, design, block=plot)
#fit a linear model for each gene
fit <- lmFit(list2,design,block=plot,correlation=cor$consensus)
cont.matrix <- makeContrasts(august=timeAugust-timeJune, levels=colnames(design))
fit.cont <- contrasts.fit(fit, cont.matrix)
fit.cont <- eBayes(fit.cont)
top.table <- topTable(fit.cont, sort.by = "none", number=Inf)
#Calculate the t statistics for each gene in each comparison
#Get the number of up and down regulated genes
#lfc 0.1=1 fold, 0.6=1.5 fold, 1=2fold
#1 fold difference
de1<-(decideTests(fit.cont,adjust.method="fdr", p.value = 0.05, lfc=0.1))
sum<-summary(de1)
sum
#1.5 fold difference
de2<-(decideTests(fit.cont,adjust.method="fdr",p.value = 0.05, lfc=0.6))
sum<-summary(de2)
sum
#2 fold difference
de3<-(decideTests(fit.cont,adjust.method="fdr",p.value = 0.05, lfc=1))
sum<-summary(de3)
sum
###Select the specific differentially expressed genes
#2 fold difference
selected3 <- top.table$adj.P.Val<0.05 & abs(top.table$logFC) >= 1
selected3 <- top.table[selected3, ]
selected3
nrow(selected3)
#save as both csv and txt file
write.csv(selected3, file = "selected_genes_fold2.csv")
write.table(selected3, file = "selected_genes_fold2.txt", quote=F, sep='          ')
#extract up and down regulated genes for 2 fold difference
UpGenes <- selected3[ selected3$logFC > 1, ]
nrow(UpGenes)
write.table(UpGenes, file = "UpAugustVsJune_2fold.txt", quote=F, sep='         ')
downGenes <- selected3[ selected3$logFC < 1, ]
nrow(downGenes)
write.table(downGenes, file = "downAugustVsJune_2fold.txt", quote=F, sep='')
# Annotation for 2 fold difference
DEG_2fold <- row.names(selected3)
DEG_2fold
write.table(DEG_2fold, file = "DEG_2fold_cwSU.csv", sep="\t", row.names = FALSE)
# Extract annotation from annotation file
file1<- read.csv("New_SA_transcriptome_Trinotate.csv", stringsAsFactors = FALSE, na.strings = "NA")
head(file1,4)
class(file1)
file2 <- read.csv("DEG_2fold_cwSU.csv", stringsAsFactors = FALSE, na.strings = "NA")
head(file2, 4)
colnames(file2)[1] <- 'transcript_id'
result = file1[file1$transcript_id %in% file2$transcript_id,]
result
head(result, 4)
```

```
#remove duplicate rows
deduped.data <- unique(result[ , 1:3 ] )
deduped.data
head(deduped.data, 4)
#Final annotation file
write.csv(deduped.data, file = "final_DEG_2fold_cwSU.csv")
##Generate heatmap
#Heatmap for 2 fold difference
names <- row.names (selected3)
highly_variable_lcpm <- logcounts[names,]
col.cell <- c("purple","orange")[time]
heatmap.2(highly_variable_lcpm,col=rev(morecols(50)),trace="none", main="Top genes across
samples",ColSideColors=col.cell,scale="row")
#Run Goseq (GO Enrichment analysis)
#for 2 fold difference
# for Upregulated genes
factor_labeling = read.table("UpAugustVsJune_2fold.txt", row.names=1)
class(factor_labeling)
head(factor_labeling,4)
factor_labeling[,1] = rep('custom_list', dim(factor_labeling)[1])
factor_labeling = factor_labeling[,1,drop=F]
colnames(factor_labeling) = c('type')
factor_list = unique(factor_labeling[,1])
DE_genes = rownames(factor_labeling)
DE_genes
# get gene lengths
gene_lengths = read.table("transcripts.length.txt", header=T, row.names=1, com='')
gene_lengths = as.matrix(gene_lengths[,1,drop=F])
gene_lengths
nrow(gene_lengths)
# get background gene list
background = read.table("background_genes.txt", header=T, row.names=1)
head(background,4)
background.gene_ids = rownames(background)
background.gene_ids
background.gene_ids = unique(c(background.gene_ids, DE_genes))
background.gene_ids
sample_set_gene_ids = background.gene_ids
sample_set_gene_ids
# parse GO assignments
GO_info = read.table("SA_go_annotations.txt", header=F, row.names=1,stringsAsFactors=F)
GO_info_listed = apply(GO_info, 1, function(x) unlist(strsplit(x,',')))
names(GO_info_listed) = rownames(GO_info)
get_GO_term_descr =  function(x) {
 d = 'none';
 go_info = GOTERM[[x]];
 if (length(go_info) >0) { d = paste(Ontology(go_info), Term(go_info), sep=' ');}
 return(d);
}
#organize go_id -> list of genes
GO_to_gene_list = list()
for (gene_id in intersect(names(GO_info_listed), sample_set_gene_ids)) {
 go_list = GO_info_listed[[gene_id]]
 for (go_id in go_list) {
   GO_to_gene_list[[go_id]] = c(GO_to_gene_list[[go_id]], gene_id)
 }
}
# GO-Seq protocol: build pwf based on ALL DE features
missing_gene_lengths = sample_set_gene_ids[! sample_set_gene_ids %in% rownames(gene_lengths)]
if (length(missing_gene_lengths) > 0) {
 stop("Error, missing gene lengths for features: ", paste(missing_gene_lengths, collapse=', '))
}
sample_set_gene_lengths = gene_lengths[sample_set_gene_ids,]
sample_set_gene_lengths
GO_info_listed = GO_info_listed[ names(GO_info_listed) %in% sample_set_gene_ids ]
```

```
cat_genes_vec = as.integer(sample_set_gene_ids %in% rownames(factor_labeling))
cat_genes_vec
pwf=nullp(cat_genes_vec, bias.data=sample_set_gene_lengths)
rownames(pwf) = sample_set_gene_ids
# perform functional enrichment testing for each category.
for (feature_cat in factor_list) {
 message('Processing category: ', feature_cat)
  gene_ids_in_feature_cat = rownames(factor_labeling)[factor_labeling$type == feature_cat]
  cat_genes_vec = as.integer(sample_set_gene_ids %in% gene_ids_in_feature_cat)
  pwf$DEgenes = cat_genes_vec
  res = goseq(pwf,gene2cat=GO_info_listed, use_genes_without_cat=TRUE)
  ## over-represented categories:
  pvals = res$over_represented_pvalue
  pvals[pvals > 1 - 1e-10] = 1 - 1e-10
  q = qvalue(pvals)
  res$over_represented_FDR = q$qvalues
  go_enrich_filename = paste("UpAugustVsJune_2fold.txt", '.GOseq.enriched', sep='')
  result_table = res[res$over_represented_pvalue<=0.05,]
  descr = unlist(lapply(result_table$category, get_GO_term_descr))
  result_table$go_term = descr;
  result_table$gene_ids = do.call(rbind, lapply(result_table$category, function(x) {
    gene_list = GO_to_gene_list[[x]]
    gene_list = gene_list[gene_list %in% gene_ids_in_feature_cat]
    paste(gene_list, collapse=', ');
  }) )
  write.table(result_table[order(result_table$over_represented_pvalue),], file=go_enrich_filename, sep='         ', quote=F,
row.names=F)
  ## under-represented categories:
  pvals = res$under_represented_pvalue
  pvals[pvals>1-1e-10] = 1 - 1e-10
  q = qvalue(pvals)
  res$under_represented_FDR = q$qvalues
  go_depleted_filename = paste("UpAugustVsJune_2fold.txt", '.GOseq.depleted', sep='')
  result_table = res[res$under_represented_pvalue<=0.05,]
  descr = unlist(lapply(result_table$category, get_GO_term_descr))
  result_table$go_term = descr;
  write.table(result_table[order(result_table$under_represented_pvalue),], file=go_depleted_filename, sep='   ', quote=F,
row.names=F)
}
# for downregulated genes
factor_labeling = read.table("downAugustVsJune_2fold.txt", row.names=1)
class(factor_labeling)
head(factor_labeling,4)
factor_labeling[,1] = rep('custom_list', dim(factor_labeling)[1])
factor_labeling = factor_labeling[,1,drop=F]
colnames(factor_labeling) = c('type')
factor_list = unique(factor_labeling[,1])
DE_genes = rownames(factor_labeling)
DE_genes
# get gene lengths
gene_lengths = read.table("transcripts.length.txt", header=T, row.names=1, com='')
gene_lengths = as.matrix(gene_lengths[,1,drop=F])
gene_lengths
nrow(gene_lengths)
# get background gene list
background = read.table("background_genes.txt", header=T, row.names=1)
head(background,4)
background.gene_ids = rownames(background)
background.gene_ids
background.gene_ids = unique(c(background.gene_ids, DE_genes))
background.gene_ids
sample_set_gene_ids = background.gene_ids
sample_set_gene_ids
# parse GO assignments
GO_info = read.table("SA_go_annotations.txt", header=F, row.names=1,stringsAsFactors=F)
```

```
GO_info_listed = apply(GO_info, 1, function(x) unlist(strsplit(x,',')))
names(GO_info_listed) = rownames(GO_info)
get_GO_term_descr =  function(x) {
 d = 'none';
 go_info = GOTERM[[x]];
 if (length(go_info) >0) { d = paste(Ontology(go_info), Term(go_info), sep=' ');}
 return(d);
}
#organize go_id -> list of genes
GO_to_gene_list = list()
for (gene_id in intersect(names(GO_info_listed), sample_set_gene_ids)) {
 go_list = GO_info_listed[[gene_id]]
 for (go_id in go_list) {
   GO_to_gene_list[[go_id]] = c(GO_to_gene_list[[go_id]], gene_id)
 }
}
# GO-Seq protocol: build pwf based on ALL DE features
missing_gene_lengths = sample_set_gene_ids[! sample_set_gene_ids %in% rownames(gene_lengths)]
if (length(missing_gene_lengths) > 0) {
 stop("Error, missing gene lengths for features: ", paste(missing_gene_lengths, collapse=', '))
}
sample_set_gene_lengths = gene_lengths[sample_set_gene_ids,]
sample_set_gene_lengths
GO_info_listed = GO_info_listed[ names(GO_info_listed) %in% sample_set_gene_ids ]
cat_genes_vec = as.integer(sample_set_gene_ids %in% rownames(factor_labeling))
cat_genes_vec
pwf=nullp(cat_genes_vec, bias.data=sample_set_gene_lengths)
rownames(pwf) = sample_set_gene_ids
# perform functional enrichment testing for each category.
for (feature_cat in factor_list) {
 message('Processing category: ', feature_cat)
 gene_ids_in_feature_cat = rownames(factor_labeling)[factor_labeling$type == feature_cat]
 cat_genes_vec = as.integer(sample_set_gene_ids %in% gene_ids_in_feature_cat)
 pwf$DEgenes = cat_genes_vec
 res = goseq(pwf,gene2cat=GO_info_listed, use_genes_without_cat=TRUE)
 ## over-represented categories:
 pvals = res$over_represented_pvalue
 pvals[pvals > 1 - 1e-10] = 1 - 1e-10
 q = qvalue(pvals)
 res$over_represented_FDR = q$qvalues
 go_enrich_filename = paste("downAugustVsJune_2fold.txt", '.GOseq.enriched', sep='')
 result_table = res[res$over_represented_pvalue<=0.05,]
 descr = unlist(lapply(result_table$category, get_GO_term_descr))
 result_table$go_term = descr;
 result_table$gene_ids = do.call(rbind, lapply(result_table$category, function(x) {
  gene_list = GO_to_gene_list[[x]]
  gene_list = gene_list[gene_list %in% gene_ids_in_feature_cat]
  paste(gene_list, collapse=', ');
 }) )
 write.table(result_table[order(result_table$over_represented_pvalue),], file=go_enrich_filename, sep='         ', quote=F,
row.names=F)
 ## under-represented categories:
 pvals = res$under_represented_pvalue
 pvals[pvals>1-1e-10] = 1 - 1e-10
 q = qvalue(pvals)
 res$under_represented_FDR = q$qvalues
 go_depleted_filename = paste("downAugustVsJune_2fold.txt", '.GOseq.depleted', sep='')
 result_table = res[res$under_represented_pvalue<=0.05,]
 descr = unlist(lapply(result_table$category, get_GO_term_descr))
 result_table$go_term = descr;
 write.table(result_table[order(result_table$under_represented_pvalue),], file=go_depleted_filename, sep='   ', quote=F,
row.names=F)
}
```